الإحصاء التطبيقي



إعداد

أ. إياد محمد الهوبي

راجعه

أ. حسام عثمان السيد

الطبعة الأولى

2017م

الفهرس

مقدمة

الحمد لله الواحد الأحد، الحمد لله الفرد الصمد، الحمد لله الذي لم يلد و لم يولد ولم يكن له كفواً أحد. والصلاة والسلام على نبينا وحبيبنا و معلمنا محمد رسول الله وعلى آله الطاهرين وصحبة الكرام الغر الميامين، ثم أما بعد.

يسعدني أن أقدم الطبعة الأولى من كتاب الإحصاء التطبيقي إلى أبنائي طلبة قسم العلوم الإدارية تخصص محاسبة وإلى الطلبة في كل مكان، آملاً أن أكون موفقاً في عرض هذه المادة العلمية بأسلوب سهل ووضوح تام.

تضمن الكتاب ستة فصول، حيث تم تخصيص الفصل الأول لمراجعة سريعة لبعض المفاهيم التي درسها الطالب في مساق مبادئ الاحصاء، الفصل الثاني عن نظرية المعاينة وتم ذكر توزيعات احتمالية جديدة لم يدرسها الطالب في مساق مبادئ الإحصاء وهي توزيع كاي تربيع وتوزيع t وتوزيع F، الفصل الثالث يتحدث عن نظرية التقدير ورَكَّزْت فيه على التقدير بفترة وتحدثت باختصار عن مفهوم المقدر وخواص المقدر الجيد، الفصل الرابع يتحدث عن اختبارات الفرضيات الخاصة بالأوساط والنسب والتباين، الفصل الخامس يتناول موضوع تحليل التباين الاحادي والثنائي وبعض اختبارات المقارنات المتعددة مثل اختبار أقل فرق معنوي واختبار دونيت، الفصل السادس يتحدث عن الاختبارات اللامعلمية. وحرصت على توحيد آلية اجراء كل الاختبارات حتى يستوعب الطالب تطبيق الاختبارات بشكل جيد وكذلك على تفسير نتائج الاختبارات ليتمكن الطالب من تطبيق ما يتعلمه نظرياً في التطبيق العملي باستخدام برنامج Spss الاحصائي، وحرصت في نهاية كل وحدة على كتابة بعض التمارين والتي من خلالها يتم تثبيت المفاهيم وتطبيق القوانين.

أقدم خالص شكري للأخ والزميل **أ. حسام عثمان السيد** لما قدمه لي من أمثلة لإثراء الكتاب ومراجعته القيمة و المفيدة.

أتمنى لجميع طلابنا وطالباتنا التفوق والنجاح سائلاً المولى عز و جل أن يهديهم سواء السبيل، إنه ولي ذلك و القادر عليه.

أ. إياد محمد الهوبي

رمضـــــــــان 1435هـ

الموافق حزيران 2014م

الفصل الأول

الاقترانات الاحتمالية

Probability functions

سنقدم في هذا الفصل مراجعة لبعض المفاهيم والنظريات التي درسها الطالب في مساق مبادئ الاحصاء وبعض المفاهيم الجديدة دون الخوض في تفاصيلها حتى يتمكن الطالب من استيعاب وفهم موضوعات الكتاب المختلفة.

المتغير العشوائي Random variable

عند إجراء تجربة عشوائية، قد يكون اهتمامنا منصباً على صفة معينة نرصدها عند إجراء التجربة العشوائية، فمثلاً عند إلقاء قطعة نقد عدة مرات قد يكون اهتمامنا منصباً على عدد الصور التي تظهر و بالتالي لا نهتم بالنتيجة التي حصلنا عليها و إنما بعدد الصور التي حصلنا عليها، هذه القيم العددية يعبر عنها بمفهوم المتغير العشوائي.

المتغير العشوائي: هو اقتران عددي معرف على فضاء العينة Ω و مجاله المقابل مجموعة الاعداد الحقيقية R ويرمز له بأحد الأحرف الإنجليزية مثل Z, Y, X وهكذا ... بمعنى:

$$\boldsymbol{X}: \boldsymbol{\Omega} \rightarrow \boldsymbol{R}$$

تصنيف المتغيرات العشوائية:

1- **متغير عشوائي منفصل Discrete random variable** حيث يكون مداه مجموعة قابلة للعد مثل عدد الطلبة في شعبة معينة، عدد المرضى في مشفى معين، عدد الحوادث في اليوم خلال أسبوع، عدد الصور التي نحصل عليها عند رمي قطعة نقد عدة مرات.

2- **متغير عشوائي متصل Continuous random variable** حيث يكون مداه فترة من الاعداد الحقيقية مثل قياس الرطوبة و درجة الحرارة و الطول و الوزن.

مثال1.1: في تجربة إلقاء قطعة نقود متزنة مرتين، أوجد مجال و مدى المتغير العشوائي X الذي يُعَبِر عن عدد الصور التي تظهر.

الحل/

مجال المتغير العشوائي هو فضاء العينة $\boldsymbol{\Omega} = \{\mathbf{HH}, \mathbf{HT}, \mathbf{TH}, \mathbf{TT}\}$

وبالتالي يكون مدى المتغير العشوائي هو $X(\boldsymbol{\Omega}) = \{\mathbf{0}, \mathbf{1}, \mathbf{2}\}$

مثال1.2: بفرض أن درجة الحرارة في غزة في شهر ديسمبر تتراوح بين 16 إلى 20 درجة مئوية فيمكن أن نُعَرِّف متغير عشوائي متصل يُعَّبِر عن درجة الحرارة خلال شهر ديسمبر مداه الفترة [**20** ، **16**].

الاقتران الاحتمالي Probability function

عند تعريف أي متغير عشوائي منفصل كان أو متصل، يمكن تعريف الاقتران الاحتمالي له والذي يكون من نفس النوع، ففي حالة النوع المنفصل يطلق على الاقتران الاحتمالي اسم اقتران الكتلة الاحتمالي **probability mass function** أو اختصاراً **(pmf)** وفي حالة النوع المتصل يطلق عليه اسم اقتران الكثافة الاحتمالي **Probability density function** أو اختصاراً **(pdf)** وهذا الاقتران يستخدم في حساب الاحتمالات للمتغير العشوائي .

تعريف:

إذا كان X متغير عشوائي منفصل معرف على فضاء عينة $\boldsymbol{\Omega}$ أي مداه $\boldsymbol{X}(\boldsymbol{\Omega})$ فإننا نعرف الاقتران الاحتمالي له ونرمز له بالرمز $\boldsymbol{p}$ كالتالي:

$$\boldsymbol{p}: \boldsymbol{X}(\boldsymbol{\Omega}) \rightarrow [\mathbf{0}, \mathbf{1}]$$

مثال1.3: أوجد الاقتران الاحتمالي للمتغير العشوائي في مثال1.1

الحل/

لاحظ أن:

$\boldsymbol{p(\{HH\}) = p(2) = 0.25}$

$\boldsymbol{p(\{HT, TH\}) = p(1) = 0.5}$

$\boldsymbol{p(\{TT\}) = p(0) = 0.25}$

ويمكن التعبير عن الاقتران الاحتمالي بجدول كالتالي:

x	0	1	2
$p(x)$	0.25	0.5	0.25

ملاحظة/ اقتران الاحتمال p للمتغير العشوائي المنفصل X يحقق ما يلي:

1) $\boldsymbol{p(x) \geq 0, \quad \forall x \in X(\Omega)}$
2) $\boldsymbol{\sum_x p(x) = 1}$

تعريف:

إذا كان X متغير عشوائي متصل معرف على فضاء عينة $\boldsymbol{\Omega}$ و مداه $[\boldsymbol{a}, \boldsymbol{b}]$ فإننا نعرف اقتران الكثافة الاحتمالي له ونرمز له بالرمز $\boldsymbol{f}$ كالتالي:

$$\boldsymbol{f: [a, b] \rightarrow [0, 1]}$$

ملاحظة/ اقتران الكثافة الاحتمالي $\boldsymbol{f}$ للمتغير العشوائي المتصل X يحقق ما يلي:

1) $\boldsymbol{f(x) \geq 0, \quad \forall x \in X(\Omega)}$
2) $\boldsymbol{\int_a^b f(x)dx = 1}$

مثال1.4: إذا كان X متغير عشوائي له اقتران كثافة احتمالي يعطى بالعلاقة

$$f(x) = \begin{cases} \frac{1}{9}x^2, & 0 < x < 3 \\ 0, & \text{elsewhere} \end{cases}$$

أوجد $P(1 < x < 4)$.

الحل/ مباشرة من تعريف الاقتران الاحتمالي نجد أن

$$P(1 < x < 4) = \int_1^4 \frac{1}{9}x^2\,dx = \int_1^3 \frac{1}{9}x^2\,dx + \int_3^\infty \frac{1}{9}x^2\,dx$$

$$= \left.\frac{x^3}{27}\right|_1^3 + 0 = \frac{27}{27} - \frac{1}{27} = \frac{26}{27}$$

مثال1.5: إذا كان X متغير عشوائي له اقتران كثافة احتمالي يعطى بالعلاقة

$$f(x) = \begin{cases} kx, & 0 < x < 2 \\ 0, & \text{elsewhere} \end{cases}$$

أوجد قيمة k ثم احسب $P(0.5 < x < 2)$.

الحل/ من الخاصية الثانية لاقتران الكثافة الاحتمالي نجد أن:

$$\int_0^2 kx\,dx = \left.\frac{kx^2}{2}\right|_0^2 = 1 \Rightarrow k = 0.5$$

فيصبح التوزيع: $f(x) = 0.5x, \quad 0 < x < 2$

وبالتالي:

$$\int_{0.5}^2 0.5x\,dx = \left.\frac{0.5x^2}{2}\right|_{0.5}^2 = \frac{15}{16}$$

اقتران التوزيع التراكمي Cumulative distribution function (cdf)

يعرف هذا الاقتران بأنه اقتران الاحتمال المتراكم لقيم المتغير العشوائي، فإذا كان X متغير عشوائي له اقتران احتمالي f فإننا نرمز للاقتران التراكمي له بالرمز F ونعرفه كالتالي:

$$\boldsymbol{F(x) = P(X \leq x)}$$

وفي حالة المتغير العشوائي المنفصل:

$$\boldsymbol{F(x) = P(X \leq x) = \sum_{-\infty}^{x} P(x)}$$

وفي حالة المتغير العشوائي المتصل:

$$\boldsymbol{F(x) = (X \leq x) = \int_{-\infty}^{x} f(x)\,dx}$$

ملاحظة/ اقتران التوزيع التراكمي يحقق ما يلي:

$\boldsymbol{1-F(-\infty) = 0}$
$\boldsymbol{2-F(\infty) = 1}$
$\boldsymbol{3-P(a \leq x \leq b) = F(x = b) - F(x = a) = F(b) - F(a)}$

مثال1.6:إذا كان X متغير عشوائي معطى اقترانه الاحتمالي بالجدول التالي:

x	5	6	7	8
$p(x)$	0.1	0.2	0.3	0.4

اوجد اقتران التوزيع التراكمي له ثم احسب $P(x \leq 4)$ ، $P(x \leq 7)$

الحل/

حيث أن:

$$P(x < 5) = 0, \;\; P(x \leq 5) = 0.1, \;\; P(x \leq 6) = 0.3$$

$$P(x \leq 7) = 0.6, \;\; P(x \leq 8) = 1$$

إذن يكون التوزيع التراكمي هو

x	5	6	7	8
$F(x)$	0.1	0.3	0.6	1

و بالتالي:

$$F(4) = P(x \leq 4) = 0$$

$$P(x \leq 7) = P(-\infty \leq x \leq 7) = F(7) - F(-\infty) = F(7) = 0.6$$

ويمكن إيجاد الاحتمال الاخير من الاقتران الاحتمالي مباشرة حيث:

$$P(x \leq 7) = P(x = 7) + P(x = 6) + P(x = 5)$$

$$= 0.3 + 0.2 + 0.1 = 0.6$$

مثال1.7: بفرض أن اقتران الكثافة الاحتمالي للمتغير العشوائي X معطى بالعلاقة

$$f(x) = \begin{cases} \frac{x^2}{9}, & 0 \leq x \leq 3 \\ 0, & \text{elsewhere} \end{cases}$$

أوجد التوزيع التراكمي ثم احسب $P(1 \leq x \leq 2)$.

الحل/

$$F(x) = \int_{-\infty}^{x} f(x)\,dx = \int_{0}^{x} \frac{t^2}{9}\,dt = \left.\frac{t^3}{27}\right|_0^x = \frac{x^3}{27}$$

$$P(1 \leq x \leq 2) = F(2) - F(1) = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$$

ويمكن إيجاد الاحتمال باستخدام اقتران الكثافة الاحتمالي كالتالي:

$$P(1 \le x \le 2) = \int_1^2 \frac{x^2}{9} dx = \left.\frac{x^3}{27}\right|_1^2 = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$$

التوقع الرياضي Mathematical expectation

قبل تعريف التوقع الرياضي سنعرف العزم الرائي (r^{th} $moment$) حول الصفر لمتغير عشوائي X.

تعريف: يعرف العزم الرائي لمتغير عشوائي X حول الصفر و يرمز له بالرمز $E(X^r)$ كما يلي:

للمتغير المنفصل : $\boldsymbol{E(X^r) = \sum_x x^r\, p_r(x)}$

للمتغير المتصل الذي مداه $[a, b]$: $\boldsymbol{E(X^r) = \int_a^b x^r\, f(x) dx}$

تعريف: ليكن X متغير عشوائي منفصل أو متصل، يعرف التوقع الرياضي أو الوسط الحسابي **Mean** للمتغير العشوائي بأنه العزم الأول حول الصفر أي $E(X)$ ، ويرمز له بالرمز μ_X ففي حالة كون المتغير منفصل يكون: $\boldsymbol{\mu_X = E(X) = \sum_x x\, P(x)}$

وفي حالة المتغير المتصل الذي مداه $[a, b]$ يكون: $\boldsymbol{\mu_X = E(X) = \int_a^b x\, f(x) dx}$

ويمثل التوقع الرياضي القيمة المتوسطة التي يأخذها المتغير العشوائي أي القيمة المتوسطة للمدى، و لاحظ أنه في حالة المتغير المنفصل إذا كانت الاحتمالات متساوية أي:

$$\boldsymbol{P(x) = \frac{1}{n}}, \qquad \forall \boldsymbol{x} \in \{\boldsymbol{x_1, x_2, x_3, \dots\dots\dots, x_n}\}$$

فإن:

$$\boldsymbol{\mu_X = E(X) = \sum_{i=1}^{n} x \frac{1}{n}} = \frac{\boldsymbol{\sum_{i=1}^{n} x}}{n}$$

وهو تعريف الوسط الحسابي للبيانات الأولية.

مثال1.8: إذا كان X متغير عشوائي له جدول التوزيع الاحتمالي التالي، أوجد توقع X.

x	0	1
$p(x)$	0.6	0.4

الحل/

$$\mu_X = E(X) = \sum_x x\,P(x) = 0 \times 0.6 + 1 \times 0.4 = 0.4$$

مثال1.9: اوجد توقع المتغير العشوائي X إذا كان اقتران كثافته الاحتمالي معطى بالعلاقة

$$f(x) = \begin{cases} x, & 0 < x < \sqrt{2} \\ 0, & \text{elsewhere} \end{cases}$$

الحل/

$$\mu_X = E(X) = \int_0^{\sqrt{2}} x\,f(x)dx = \int_0^{\sqrt{2}} x\,.\,xdx$$

$$= \int_0^{\sqrt{2}} x^2\,dx = \left.\frac{x^3}{3}\right|_0^{\sqrt{2}} = \frac{2\sqrt{2}}{3}$$

تعريف: يعرف تباين **(Variance)** المتغير العشوائيX ويرمز له بالرمز $\boldsymbol{\sigma_x^2}$ أو بالرمز $\boldsymbol{Var(X)}$ بأنه العزم الثاني حول الوسط أي:

$$\boldsymbol{Var(X) = \sigma_X^2 = \sum_x (x - \mu_X)^2}$$

أو بصورة مكافئة:

$$\boldsymbol{Var(X) = \sigma_X^2 = E(X^2) - (\mu_X)^2}$$

تعريف: يعرف الانحراف المعياري لمتغير عشوائي X (منفصل أو متصل) ويرمز له بالرمز σ_X بأنه الجذر التربيعي للتباين أي:

$$\sigma_X = \sqrt{\sigma_X^2}$$

ملاحظة/ إذا كان X متغير عشوائي (منفصل أو متصل) وكان $Y = g(X)$ فإن Y متغير عشوائي من نفس نوع X ويكون توقع Y كما يلي:

في حالة كون المتغير منفصل يكون: $E(g(X)) = \sum_x g(x)\, P(x)$

وفي حالة المتغير المتصل الذي مداه $[a, b]$ يكون: $E(g(X)) = \int_a^b g(x)\, f(x) dx$

الآن سنذكر بعض النظريات الهامة في التوقع بدون برهان.

نظرية 1.1: إذا كان X متغير عشوائي، c ثابت اختياري فإن:

$$E(c) = c$$

$$E(cX) = cE(X)$$

نظرية1.2: إذا كان Y, X متغيران عشوائيان، فإن:

$$E(X \pm Y) = E(X) \pm E(Y).$$

نظرية 1.3: إذا كان Y, X متغيران عشوائيان مستقلان، فإن :

$$E(XY) = E(X)E(Y)$$

نظرية 1.4: إذا كان X متغير عشوائي، c ثابت اختياري فإن:

$$Var(c) = 0$$

$$Var(cX) = c^2 Var(X)$$

نظرية1.5: إذا كان Y, X متغيران عشوائيان مستقلان، فإن :

$$\boldsymbol{Var(X \pm Y) = Var(X) + Var(Y)}$$

أو بصورة مكافئة :

$$\boldsymbol{\sigma^2_{X\pm Y} = \sigma^2_X + \sigma^2_Y}$$

مثال1.10: ليكن X متغير عشوائي يعطى اقترانه الاحتمالي بالجدول:

x	0	1	2	3
$p(x)$	0.2	0.3	0.4	0.1

أوجد:. $E(X)$, σ_X^2 , σ_X , $E(X^2+2X)$

الحل/

$$E(X) = \sum_x x\,P(x) = 0\times 0.2 + 1\times 0.3 + 2\times 0.4 + 3\times 0.1 = 1.4$$

$$E(X^2) = \sum_x x^2\,P(x) = 0\times 0.2 + 1\times 0.3 + 4\times 0.4 + 9\times 0.1 = 2.8$$

$$\sigma_X^2 = E(X^2) - (\mu_X)^2 = 2.8 - (1.4)^2 = 0.84$$

$$\sigma_X = \sqrt{\sigma_X^2} = \sqrt{0.84} = 0.92$$

$$E(X^2+2X) = \sum_x (x^2+2x)\,P(x)$$

$$= 0\times 0.2 + 3\times 0.3 + 8\times 0.4 + 15\times 0.1 = 5.6$$

مثال1.11:أوجد التوقع و الانحراف المعياري للمتغير العشوائي X إذا كان اقتران الكثافة الاحتمالي له معطى بالعلاقة

$$f(x) = \begin{cases} \frac{x}{2}, & 0 < x < 2, \\ 0, & \text{elsewhere} \end{cases}$$

ثم اوجد $E(X+1)$

الحل/

$$E(X) = \int_0^2 x.\frac{x}{2}\,dx = \int_0^2 \frac{x^2}{2}\,dx = \frac{x^3}{6}\Big|_0^2 = \frac{8}{6}$$

$$E(X^2) = \int_0^2 x^2.\frac{x}{2}\,dx = \int_0^2 \frac{x^3}{2}\,dx = \frac{x^4}{8}\Big|_0^2 = \frac{16}{8} = 2$$

$$\sigma_X^2 = E(X^2) - (\mu_X)^2 = 2 - \left(\frac{8}{6}\right)^2 = 0.22$$

$$\sigma_X = \sqrt{\sigma_X^2} = \sqrt{0.22} = 0.47$$

$$E(X+1) = \int_0^2 (x+1).\frac{x}{2}\,dx = \int_0^2 (\frac{x^2}{2} + \frac{x}{2})\,dx = (\frac{x^3}{6} + \frac{x^2}{4})\Big|_0^2 = \frac{7}{3}$$

توزيعات احتمالية خاصة Special probability distributions

التوزيعات الاحتمالية الخاصة هي عبارة عن اقترانات احتمالية لها قاعدة معروفة، وهي تنتج عن تجارب عشوائية لوحظ أنها تأخذ نمطاً معيناً في تغيرها مكن المختصين من الحصول على صيغ رياضية تصف تغيرات نتائج تلك التجارب العشوائية.

وسنتحدث عن التوزيع الطبيعي – التوزيع المنتظم المستمر– التوزيع الأسي وتوزيع جاما كأمثلة على التوزيعات المتصلة و عن توزيع ذات الحدين و التوزيع المنتظم و توزيع بواسون كأمثلة على النوع المنفصل، حيث سنعرف الاقترانات الاحتمالية لها ونذكر بعض خواصها وبعض مؤشراتها الاحصائية.

التوزيع الطبيعي The Normal distribution

يعتبر هذا التوزيع من أهم التوزيعات الاحتمالية و أكثرها استخداما، ويلاحظ أن أغلب الظواهر الطبيعية تتبع في توزيعها التوزيع الطبيعي، مثل أعمار الكائنات الحية و أوزان المنتجات الزراعية و الصناعية.

فإذا كان X متغير عشوائي له توزيع طبيعي وسطه الحسابي $\boldsymbol{\mu}$ وتباينه $\boldsymbol{\sigma^2}$ فإن الاقتران الاحتمالي له هو:

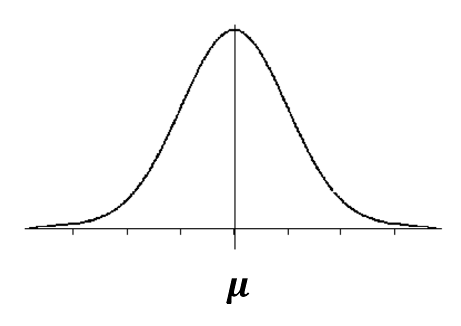
$$\boldsymbol{f(x)} = \frac{\mathbf{1}}{\boldsymbol{\sigma\sqrt{2\pi}}} \boldsymbol{e}^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \qquad , -\infty < \boldsymbol{x} < \infty$$

ونعبر عن ذلك بالصيغة

$$\boldsymbol{X} \sim \mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\sigma^2})$$

والتي تُقرأ " X تتبع التوزيع الطبيعي بوسط μ وتباين σ^2"

والرسم التالي يوضح منحنى التوزيع:



يتضح من الرسم أن التوزيع متماثل حول الوسط الحسابي وأن أكثر المشاهدات تقع حول الوسط الحسابي و أقلها في الطرفين.

في الحالة الخاصة التي يكون فيها $\boldsymbol{\mu} = \mathbf{0}$ و $\boldsymbol{\sigma^2} = \mathbf{1}$ يطلق على التوزيع اسم التوزيع الطبيعي المعياري **Standard normal distribution** ونعبر عن ذلك بالصيغة:

$$\boldsymbol{X} \sim \mathbf{N}(\mathbf{0}, \mathbf{1})$$

ويكون الاقتران الاحتمالي للتوزيع في هذه الحالة معطى بالعلاقة

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \qquad -\infty < x < \infty$$

نظرية1.6: إذا كان $X \sim N(\mu, \sigma^2)$ فإن $z = \frac{X-\mu}{\sigma} \sim N(0,1)$.

إن قيم z تسمى القيم المعيارية وهي قيم خالية من وحدات القياس ويتم حساب المساحات أسفل المنحنى باستخدام جدول (1) الخاص بالتوزيع المعياري والملحق في آخر الكتاب.

بعض المؤشرات الاحصائية للتوزيع:

Mean	μ
Variance	σ^2
Standard division	σ

التوزيع المنتظم المستمر The Continuous Uniform distribution

إذا كان X متغير عشوائي مستمر و معرفا على الفترة $[a, b]$ وله اقتران كثافة احتمالي هو:

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b, \\ 0, & \text{elsewhere} \end{cases}$$

فإن هذا التوزيع هو التوزيع المنتظم المستمر بالمعلمتين a و b

ونعبر عن ذلك بالصيغة

$$X \sim U(a, b)$$

بعض المؤشرات الاحصائية للتوزيع:

Mean	$\mu = \frac{a+b}{2}$
Variance	$\sigma^2 = \frac{(b-a)^2}{12}$
Standard division	$\sigma = \sqrt{\frac{(b-a)^2}{12}}$

مثال1.12: اتفق صديقان على أن يلتقيا في الجامعة في حدود الساعة 10−10:30 ، فما احتمال حدوث لقائهما قبل الساعة 10:15.

الحل/

من الواضح أن الزمن X متغير عشوائي مستمر، وعند تمام الساعة العاشرة يبدأ احتمال اللقاء وذلك خلال 30 دقيقة، لذلك تبدأ فترة اللقاء بالصفر وتنتهي بـ30 دقيقة، فيكون لدينا:

$$X \sim \mathrm{U}(0,30)$$

والاقتران الاحتمالي هو:

$$f(x) = \begin{cases} \frac{1}{30}, & 0 \le x \le 30 \\ 0, & \text{elsewhere} \end{cases}$$

أما الاحتمال المطلوب فهو:

$$P(x \le 15) = \int_0^{15} \frac{1}{30}\, dx = 0.5$$

والمؤشرات الاحصائية للتوزيع:

Mean	15
Variance	75
Standard division	8.66

The Exponential distribution التوزيع الأسي

في حالة دراسة الزمن الذي يمضي بين وقوع حادث ووقوع حادث آخر فإن التوزيع المناسب في هذه الحالة هو التوزيع الأسي.

فإذا كان X متغير عشوائي له توزيع أسي فإن الاقتران الاحتمالي له هو:

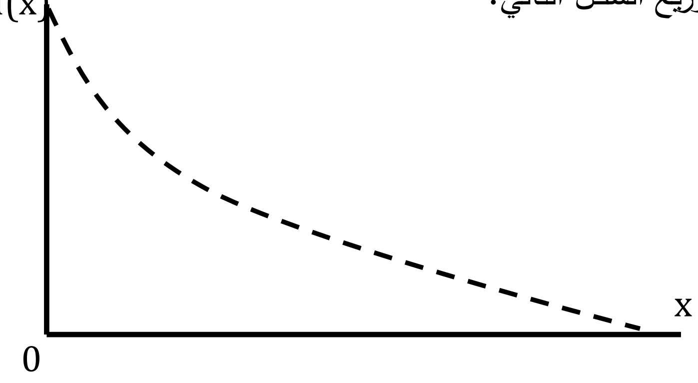
$$\boldsymbol{f(x)} = \begin{cases} \boldsymbol{ke^{-kx}}, & \boldsymbol{x > 0}, \\ \boldsymbol{0}, & \textbf{elsewhere} \end{cases}$$

حيث k ثابت يمثل معلمة التوزيع.

ونعبر عن ذلك بالصيغة

$$\boldsymbol{X \sim \mathrm{Exp}(k)}$$

ويأخذ هذا التوزيع الشكل التالي:



بعض المؤشرات الاحصائية للتوزيع:

Mean	$\boldsymbol{\mu = \frac{1}{k}}$
Variance	$\boldsymbol{\sigma^2 = \frac{1}{k^2}}$
Standard division	$\boldsymbol{\sigma = \frac{1}{k}}$

مثال1.13: في معمل للمشروبات الغازية لوحظ أن معدل زمن انتظار الحصول على قنينة غير صالحة للتسويق هو 3 ثوان، فما احتمال مرور أكثر من 4 ثانية دون ملاحظة قنينة غير صالحة و اوجد المؤشرات الاحصائية للتوزيع.

الحل/

واضح أن التوزيع هنا هو الأسي بمعدل$\mu = 3$ ، لذلك تكون معلمة التوزيع $k = \frac{1}{\mu} = \frac{1}{3}$ ويكون الاقتران الاحتمالي للتوزيع هو:

$$f(x) = \begin{cases} \frac{1}{3} e^{-\frac{1}{3}x}, & x > 0, \\ 0, & \text{elsewhere} \end{cases}$$

فيكون الاحتمال المطلوب هو:

$$P(x > 4) = \int_{4}^{\infty} \frac{1}{3} e^{-\frac{1}{3}x}\, dx = -e^{-\frac{1}{3}x}\Big|_{4}^{\infty} = 0 + e^{-\frac{4}{3}} = 0.264$$

والمؤشرات الاحصائية للتوزيع:

Mean	3
Variance	9
Standard division	3

توزيع ذي الحدين The Binomial distribution

في التجارب العشوائية التي تجرى عدة مرات مستقلة وفي كل مرة نحصل إما على نجاح أو فشل و احتمال النجاح ثابت في كل مرة، يكون عدد مرات النجاح يمثل متغيرا عشوائيا له توزيع احتمالي يسمى توزيع ذي الحدين.

بفرض X هو المتغير العشوائي الذي يعبر عن عدد مرات النجاح وكان احتمال النجاح **p** وكان **n** عدد مرات تكرار التجربة، فإن اقتران الاحتمال للمتغير العشوائي (توزيع ذي الحدين) هو:

$$\boldsymbol{P(x) = \binom{n}{x} \mathbf{p}^{x}(1-\mathbf{p})^{n-x}, \qquad x = 0, 1, 2, \dots\dots, n}$$

حيث $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ ، $\boldsymbol{x}$ تمثل عدد مرات النجاح وتعتبر $\mathbf{n, p}$ هما معلمتي هذا التوزيع.

ونعبر عن ذلك بالصيغة

$$\boldsymbol{X \sim \mathbf{b}(\mathbf{n}, \mathbf{p})}$$

بعض المؤشرات الاحصائية للتوزيع:

Mean	$\boldsymbol{\mu = \mathbf{np}}$
Variance	$\boldsymbol{\sigma^2 = \mathbf{np}(1-\mathbf{p})}$
Standard division	$\boldsymbol{\sigma = \sqrt{\mathbf{np}(1-\mathbf{p})}}$

مثال1.15: إذا كان X متغير عشوائي بحيث $X \sim \mathrm{b}(4,0.2)$، أوجد ما يلي:

$$P(x \geq 1), \quad P(x = 2), \quad P(x)$$

الحل/

من الواضح أن $\mathrm{n} = 4, \ \mathrm{p} = 0.2$ لذلك:

$$P(x) = \binom{4}{x}(0.2)^x(0.8)^{4-x}, \qquad x = 0,1,2,3,4$$

$$P(2) = \binom{4}{2}(0.2)^2(0.8)^2 = 0.1536$$

$$P(x \geq 1) = 1 - P(x = 0) = 1 - \binom{4}{0}(0.2)^0(0.8)^4 = 0.5904$$

مثال1.16: إذا كان $X \sim \text{b}(3,0.3)$ ، $Y \sim \text{b}(30,0.5)$ أوجد المؤشرات الاحصائية للتوزيعين.

الحل/

	X	Y
Mean	0.9	15
Variance	0.63	7.5
Standard division	0.79	2.74

التوزيع المنتظم The Uniform distribution

أحيانا عند اجراء تجربة عشوائية، يكون جميع النتائج لها نفس فرصة الحدوث، أي نفس الاحتمال، فنقول في هذه الحالة أن النتائج تمثل توزيعا منتظماً.

فإذا كان X متغير عشوائي له توزيع منتظم فإن الاقتران الاحتمالي له هو:

$$\boldsymbol{P(x) = \frac{1}{n}, \qquad x = 0, 1, 2, \dots \dots, n}$$

حيث $\boldsymbol{n}$ عدد موجب وهو معلمة هذا التوزيع.

ونعبر عن ذلك بالصيغة

$$X \sim \mathrm{U}(\mathrm{n})$$

والمؤشرات الاحصائية للتوزيع:

Mean	$\mu = \frac{n+1}{2}$
Variance	$\sigma^2 = \frac{\mathrm{n}^2 - 1}{12}$
Standard division	$\sigma = \sqrt{\frac{\mathrm{n}^2 - 1}{12}}$

مثال1.17: بفرض أن X متغير عشوائي له توزيع منتظم معطى بالعلاقة:

$$P(x) = \frac{1}{9}, \qquad x = 0,1,2,\dots\dots,9$$

أوجد المؤشرات الاحصائية للتوزيع.

الحل/

Mean	5
Variance	6.67
Standard division	2.58

توزيع بواسون Poisson distribution

يعرف هذا التوزيع بتوزيع الحوادث النادرة، حيث أنه يصلح للحوادث نادرة الحدوث مثل عدد حوادث سقوط الطائرات وعدد وصول رسائل بالخطأ لبريد المدينة، ويمثل هذا التوزيع حالة خاصة من توزيع ذي الحدين وذلك عندما يكون احتمال النجاح صغير جدا مقابل عدد تكرارات كبير جدا.

إذا كان X متغير عشوائي له توزيع بواسون فإن الاقتران الاحتمالي له هو:

$$\boldsymbol{P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \qquad x = 0, 1, 2, \ldots \ldots \ldots}$$

ونعبر عن ذلك بالصيغة

$$\boldsymbol{X \sim \mathrm{PO}(\lambda)}$$

والمؤشرات الاحصائية للتوزيع:

Mean	$\boldsymbol{\mu = \lambda}$
Variance	$\boldsymbol{\sigma^2 = \lambda}$
Standard division	$\boldsymbol{\sigma = \sqrt{\lambda}}$

مثال1.18: إذا كان X متغير عشوائي بحيث $X \sim \mathrm{PO}(4)$ ، احسب $P(x > 1)$ و المؤشرات الاحصائية للتوزيع.

الحل/

اقتران الاحتمال للمتغير العشوائي هو

$$P(x) = \frac{e^{-4}4^{x}}{x!}, \qquad x = 0,1,2, \ldots \ldots \ldots$$

فيكون المطلوب

$$P(x > 1) = \mathbf{1} - P(x \le 1) = 1 - \left\{\frac{e^{-4}4^{0}}{0!} + \frac{e^{-4}4^{1}}{1!}\right\} = 0.91$$

و المؤشرات الاحصائية للتوزيع:

Mean	4
Variance	4
Standard division	2

تمارين:

1- الجدول التالي يمثل اقتران احتمالي لمتغير عشوائي X و المطلوب ايجاد قيمة الثابت k ثم حساب تباين المتغير العشوائي .

x	2	4	6	8
$P(x)$	0.1	k	2k	0.3

2- إذا كان X متغير عشوائي له اقتران كثافة احتمالية $0 < x < 2$, $f(x) = \frac{x}{c}$ فما هي القيمة المتوقعة لمربعات قيم X حيث c مقدار ثابت.

3- إذا كان X متغير عشوائي له اقران احتمالي معطى بالعلاقة

$$f(x) = k(2x+1), \quad x = 0,1,\ldots\ldots,6$$

أوجد ما يلي:

1- قيمة الثابت k.

2- $F(x)$

3- $P(0 \le x \le 4)$

4- $P(x > 3)$

5- التباين

4- إذا كان X متغير عشوائي له اقتران توزيع تراكمي معطى بالعلاقة

$$F(x) = \frac{(x-1)^2}{k}, \quad 1 < x < 3$$

أوجد قيمة k ثم اوجد الاقتران الاحتمالي له.

5− إذا كان $X \sim \mathrm{b}(4,0.3)$ احسب ما يلي:

- خصائص التوزيع الاحصائية.
- $P(0 < x \le 3)$
- ارسم الاقتران الاحتمالي.

6− إذا كان $X \sim \mathrm{U}(6)$ احسب ما يلي:

- خصائص التوزيع الاحصائية.
- $P(2 < x < 7)$
- ارسم الاقتران الاحتمالي.

7- تفخر شركة أن نسبة المعيب في انتاجها اليومي الذي يبلغ 1500 قطعة كانت 0.005

أوجد ما يلي:

- اقتران التوزيع الاحتمالي للمتغير X الذي يعبر عن عدد القطع المعيبة.
- خصائص التوزيع الاحصائية.
- $P(x > 8)$
- ارسم الاقتران الاحتمالي.

8- إذا كان $X \sim \mathrm{U}(5,15)$ احسب ما يلي:

- الاقتران الاحتمالي للتوزيع.
- خصائص التوزيع الاحصائية.
- $P(2 < x < 7)$

9- إذا كان $X \sim \mathrm{Exp}(\frac{1}{5})$ اوجد:

- الاقتران الاحتمالي للتوزيع.
- خصائص التوزيع الاحصائية.

الفصل الثاني

نظرية المعاينة

Sampling theory

تحدثنا في مساق مبادئ الاحصاء عن أساليب جمع البيانات وقلنا أن هناك اسلوبين هما اسلوب المسح الشامل واسلوب المعاينة، وتحدثنا عن طرق جمع البيانات وعن بعض العينات العشوائية وكيفية سحبها من المجتمع مثل العينة العشوائية البسيطة، العينة العشوائية الطبقية، العينة العشوائية متعددة المراحل و العينة العشوائية المنتظمة، وسنتحدث في هذا الفصل عن توزيعات خاصة تسمى توزيعات المعاينة، حيث تلعب هذه التوزيعات دوراً رئيساً في اختبار الفرضيات وفي الاحصاء التطبيقي بشكل عام.

نعرف في البداية مفهوم المعاينة، فالمعاينة هي عملية الحصول على عينة من المجتمع، فلو أردنا سحب عينة من المجتمع حجمها $\boldsymbol{n}$ و كان حجم المجتمع $\boldsymbol{N}$ وكان السحب مع الارجاع فعدد الطرق الممكنة لسحب العينة هو $\boldsymbol{N^n}$ ، أما لو كان السحب بدون ارجاع فعدد الطرق الممكنة هو $\binom{\mathbf{N}}{\mathbf{n}}$ حيث $\binom{\mathbf{N}}{\mathbf{n}} = \frac{\boldsymbol{N}!}{\boldsymbol{n}!(\boldsymbol{N}-\boldsymbol{N})!}$ ، فمثلاً لو كان حجم المجتمع 10 و حجم العينة 3 و كان السحب مع الارجاع فيمكن سحب $10^3 = 1000$ عينة مختلفة، أما إذا كان السحب بدون ارجاع فيمكن سحب $\binom{\mathbf{10}}{\mathbf{3}} = 120$ عينة مختلفة.

ذكرنا في الفصل الأول أهمية التوزيع الطبيعي في التطبيقات الاحصائية الواسعة، وسنذكر الآن بدون برهان نظرية هامة تسمى نظرية النهاية المركزية والتي تلعب دوراً هاماً في الاحصاء التطبيقي حيث تمكننا من تقريب أي توزيع احتمالي للتوزيع الطبيعي عند تحقق شروط معينة.

نظرية 2.1: ليكن $\{X_i\}_{i=1}^{n}$ مجموعة من المتغيرات العشوائية المستقلة والتي تتبع نفس التوزيع وكلٍ منها بوسط μ وتباين σ^2. إذا كان $S_n = \sum X_i$ فإن:

$$\frac{\boldsymbol{S_n - n\mu}}{\boldsymbol{\sigma\sqrt{n}}} \sim \mathbf{N}(\mathbf{0}, \mathbf{1})$$

إن للتوزيع الطبيعي دور مهم آخر يتمثل في اشتقاق توزيعات احتمالية متعلقة بالمعاينة وذلك عند سحب عينة صغيرة حجمها $(n < 30)$ تسحب عادة من مجتمع طبيعي وهذه التوزيعات هي:

1- توزيع مربع كاي χ^2 .

2- توزيع t .

3- توزيع F .

و سنتحدث عن كل توزيع بشيء من التفصيل.

توزيع مربع كاي χ^2 The chi-Square distribution

نعلم من مساق مبادئ الاحصاء أنه إذا كان X متغير عشوائي بحيث $X \sim N(\mu, \sigma^2)$ فإن:

$$z = \frac{X-\mu}{\sigma} \sim N(0,1)$$

وعليه فإننا نُعَرِف التوزيع الاحتمالي لمربعات قيم z بأنه توزيع مربع كاي بدرجة حرية $\vartheta = 1$ ويرمز له بالرمز χ^2، حيث درجة الحرية هي معلمة هذا التوزيع. ونعبر عن ذلك بالصيغة:

$$z^2 = \left(\frac{X-\mu}{\sigma}\right)^2 \sim \chi^2_{(1)}$$

ويمكن ايجاد قيم متغير يتبع هذا التوزيع باستخدام الجدول الخاص بالتوزيع في نهاية الكتاب.

إن اقتران الكثافة الاحتمالي للتوزيع بدرجة حرية ϑ يعطى بالصيغة:

$$f(x) = \frac{1}{\Gamma(\frac{\vartheta}{2})(2)^{\frac{\vartheta}{2}}} x^{\frac{\vartheta}{2}-1} e^{-\frac{x}{2}}, \qquad x \geq 0$$

واضح أن هذا التوزيع هو حالة خاصة من توزيع جاما بالمعلمتين $\alpha = \frac{\vartheta}{2}$ و $\beta = 2$

سنذكر هنا بعض النظريات الهامة دون برهان.

نظرية 2.2:إذا كانت $X_1, X_2, X_3 \dots \dots \dots \dots X_n$ متغيرات عشوائية مستقلة تتبع التوزيع الطبيعي المعياري، فإن:

$$\sum_{i=1}^{n} X_i^2 \sim \chi^2_{(n)}$$

نظرية 2.3:إذا كانت $U_1, U_2, U_3 \dots \dots \dots \dots U_n$ متغيرات عشوائية مستقلة تتبع توزيع كاي تربيع بدرجات حرية $\vartheta_1, \vartheta_2, \vartheta_3 \dots \dots \vartheta_n$ على الترتيب، فإن:

$$\sum_{i=1}^{n} U_i \sim \chi^2_{(\vartheta)}$$

حيث $\vartheta = \sum_{i=1}^{n} \vartheta_i$.

نظرية 2.4:إذا اخذت عينات عشوائية كلٍ بحجم n وتباين S^2 من مجتمع طبيعي التوزيع بوسط μ وتباين σ^2، فإن:

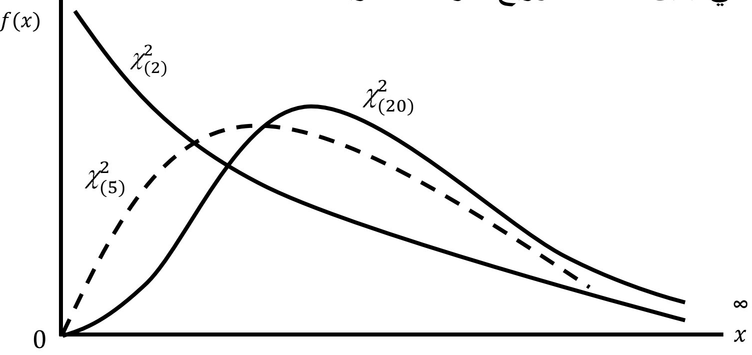
$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

حيث $S^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$

والجدول التالي يبين بعض المؤشرات الاحصائية للتوزيع:

Mean	$\mu = \vartheta$
Variance	$\sigma^2 = 2\vartheta$
Standard division	$\sigma = \sqrt{2\vartheta}$

والرسم التالي يبين شكل التوزيع لدرجات حرية مختلفة:



يتضح من الرسم أن التوزيع موجب الالتواء و أنه يقترب من التماثل كلما زادت درجة الحرية و في حالة $n \geq 30$ يمكن تقريب التوزيع للتوزيع الطبيعي.

ونعبر عن قيمة المتغير العشوائي الذي يتبع توزيع مربع كاي بالرمز $\chi^2_{(\alpha,\vartheta)}$ وهي القيمة التي يقع على يمينها مساحة α على منحنى كاي تربيع بدرجة حرية ϑ وتحسب قيم هذا المتغير العشوائي والاحتمالات له باستخدام جدول (2) المرفق في نهاية الكتاب، والمثال التالي يوضح كيفية استخدام ذلك الجدول الخاص بالتوزيع.

مثال2.1: اوجد القيم $\chi^2_{(0.1,17)}$, $\chi^2_{(0.995,8)}$, $\chi^2_{(0.95,7)}$

الحل/

لإيجاد $\chi^2_{(0.95,7)}$ نختار من العمود الايسر (عمود درجة الحرية) القيمة 7 فنجد على يمينه صف من الأعداد ومن الصف العلوي (صف المساحات أي الاحتمالات) عن القيمة 0.95 فنجد أسفلها عمود من الاعداد فتكون القيمة الموجودة في تقاطع الصف و العمود هي القيمة المطلوبة وبالتالي يكون $\chi^2_{(0.95,7)} = 2.167$ وبالمثل نجد أن:

$$\chi^2_{(0.995,8)} = 1.344 \quad , \chi^2_{(0.1,17)} = 24.769$$

والجدير بالذكر أن توزيع كاي تربيع له استخدامات كثيرة في الاحصاء التطبيقي نذكر منها:

1- اختبار قيمة تباين مجتمع طبيعي التوزيع.
2- اختبار حسن المطابقة (توفيق البيانات)
3- اختبار الاستقلالية.
4- اختبار تجانس عدة تقديرات مستقلة لتباين المجتمع الطبيعي.
5- اختبار معامل الارتباط.
6- ايجاد فترات الثقة لتباين المجتمع.

توزيع Student's distribution t

إذا كان X متغير عشوائي له اقتران كثافة احتمالي معطى بالعلاقة:

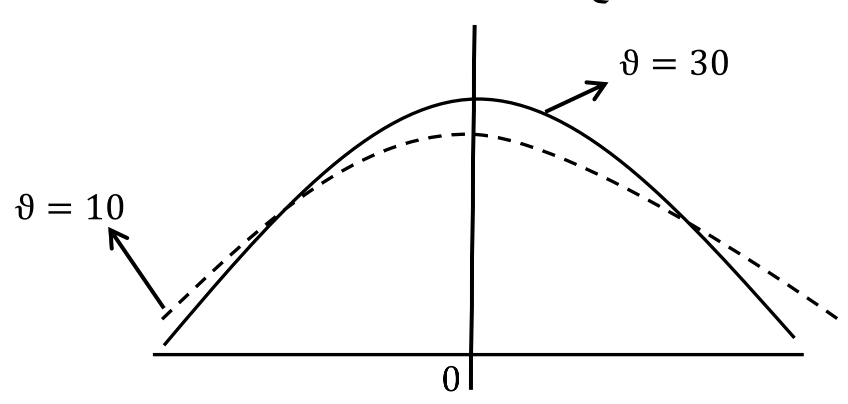
$$f(t) = \frac{\Gamma(\frac{\vartheta+1}{2})}{\Gamma(\frac{\vartheta}{2})\sqrt{\vartheta\pi}}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty < t < \infty$$

نقول أن المتغير يتبع توزيع t بدرجة حرية ϑ والتي هي بدورها معلمة هذا التوزيع، ويلعب هذا التوزيع دورا هاما عند سحب عينات صغيرة الحجم $n < 30$ من مجتمع مجهول التباين.

والجدول التالي يبين بعض المؤشرات الاحصائية للتوزيع:

Mean	$\mu = 0$
Variance	$\sigma^2 = \frac{n}{n-2}, \; n > 2$
Standard division	$\sigma = \sqrt{\frac{n}{n-2}}$

والرسم التالي يبين شكل منحنى التوزيع لدرجات حرية مختلفة:



يلاحظ من الرسم أن التوزيع متماثل دائما حول الوسط الحسابي، ويميل للاعتدال عند درجات الحرية الكبيرة $\boldsymbol{n \geq 30}$ ،

وسنذكر النظريات الهامة التالية بدون برهان والتي سنستخدمها في تطبيقاتنا.

نظرية2.5: إذا كان Y ، Z متغيران عشوائيان مستقلان بحيث $Y \sim \mathrm{N}(0,1)$ و $\mathrm{Z} \sim \chi^2_{(\vartheta)}$ فإن:

$$\boldsymbol{T = \frac{Y}{\sqrt{Z/\vartheta}} \sim t(\vartheta)}$$

نظرية2.6: إذا اخذت عينات عشوائية كلٍ بحجم n وتباين S^2 من مجتمع طبيعي التوزيع بوسط μ وتباين σ^2، فإن:

$$\mathbf{T} = \frac{\boldsymbol{\bar{X} - \mu}}{\boldsymbol{S/\sqrt{n-1}}} \sim \boldsymbol{t(n-1)}$$

و نعبر عن قيمة المتغير العشوائي الذي يتبع توزيع t بالرمز $t_{(\alpha,\vartheta)}$ وهي القيمة التي يقع على يمينها مساحة α على منحى توزيع t بدرجة حرية ϑ. وتحسب قيم هذا المتغير العشوائي والاحتمالات له باستخدام جدول (3) المرفق في نهاية الكتاب.

ملاحظة/ من تماثل التوزيع حول الصفر يكون $t_{(\alpha,\vartheta)} = -t_{(1-\alpha,\vartheta)}$

مثال2.2: اوجد القيم $t_{(0.05,3)}$، $t_{(0.99,11)}$، $t_{(0.995,9)}$

الحل/

لإيجاد $t_{(0.05,3)}$ نختار من العمود الايسر (عمود درجة الحرية) القيمة 3 فنجد على يمينه صف من الاعداد ومن الصف العلوي (صف المساحات أي الاحتمالات) عن القيمة 0.05 فنجد أسفلها عمود من الاعداد فتكون القيمة الموجودة في تقاطع الصف و العمود هي القيمة المطلوبة وبالتالي يكون $t_{(0.05,3)} = 2.353$.

لإيجاد $t_{(0.99,11)}$ نلاحظ أن جدول التوزيع لا يحتوي على القيمة $\alpha = 0.99$ لذلك نستغل تماثل منحنى التوزيع حول الصفر فيكون:

$$t_{(0.99,11)} = -t_{(0.01,11)} = -2.718$$

وبالمثل يكون $t_{(0.995,9)} = -t_{(0.005,9)} = -3.250$

إن لتوزيع t استخدامات كثيرة منها:

1- اختبار متوسط مجتمع طبيعي مجهول التباين.
2- اختبار الفرق بين متوسطي مجتمعين طبيعيين مجهولي التباين.
3- اختبار معنوية معامل الارتباط البسيط.
4- اختبار معنوية معاملات الانحدار في الانحدار الخطي المتعدد.
5- اختبار معنوية معامل الارتباط الجزئي.
6- تكوين فترات الثقة لمتوسط مجتمع طبيعي مجهول التباين.

توزيع F   F- distribution

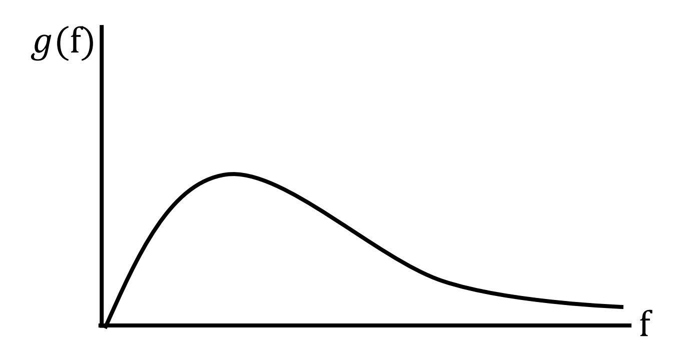
إذا كان X_2, X_1 متغيرين عشوائيين مستقلين بحيث $X_1 \sim \chi^2_{(\vartheta_1)}$ ، $X_2 \sim \chi^2_{(\vartheta_2)}$ فإن النسبة: $\mathbf{f} = \frac{X_1/\vartheta_1}{X_2/\vartheta_2}$ بين التوزيعين تعرف توزيع F بالمعلمتين ϑ_2, ϑ_1 وهما درجة حرية البسط و المقام على الترتيب، و اقتران الكثافة الاحتمالي للتوزيع معطى بالعلاقة:

$$\boldsymbol{g}(\mathbf{f}) = \frac{\Gamma(\frac{\vartheta_1 + \vartheta_2}{2})}{\Gamma(\frac{\vartheta_1}{2})\Gamma(\frac{\vartheta_2}{2})} (\vartheta_1)^{\vartheta_1/2} (\vartheta_2)^{\vartheta_2/2} \frac{\mathbf{f}^{\frac{\vartheta_1}{2}-1}}{(\vartheta_2 + \vartheta_1 \mathbf{f})^{\frac{\vartheta_1+\vartheta_2}{2}}}, \quad \mathbf{f} > 0$$

والجدول التالي يبين بعض المؤشرات الاحصائية للتوزيع:

Mean	$\boldsymbol{\mu = \frac{\vartheta_2}{\vartheta_2 - 2}, \quad \vartheta_2 > 2}$
Variance	$\boldsymbol{\sigma^2 = \frac{2{\vartheta_2}^2(\vartheta_1 + \vartheta_2 - 2)}{\vartheta_1(\vartheta_2 - 2)^2(\vartheta_2 - 4)}}$
Standard division	$\boldsymbol{\sigma = \sqrt{\frac{2{\vartheta_2}^2(\vartheta_1 + \vartheta_2 - 2)}{\vartheta_1(\vartheta_2 - 2)^2(\vartheta_2 - 4)}}}$

ويقترب شكل التوزيع من الشكل التالي:



و نعبر عن قيمة المتغير العشوائي الذي يتبع توزيع F بالرمز $F_{(\alpha,\vartheta_1,\vartheta_2)}$ وهي القيمة التي يقع على يمينها مساحة α على منحى توزيع F بدرجة حرية ϑ_1 في البسط و ϑ_2 في المقام.

و جدول (4) المرفق في نهاية الكتاب خاص لقيم α التالية:

{ 0.01 ، 0.025 ، 0.05 ، 0.1 }

مثال2.3: أوجد القيم $F_{(0.05,1,7)}$ ، $F_{(0.025,3,5)}$ ، $F_{(0.1,12,8)}$

الحل/

لحساب $F_{(0.05,1,7)}$ نختار من جداول التوزيع الجدول الخاص بقيمة $\alpha = 0.05$ ، ومن الصف في أعلى الجدول الخاص بدرجات حرية البسط نختار القيمة 1 حيث يوجد أسفلها عمود من الاعداد، ومن العمود أقصى اليسار والخاص بدرجة حرية المقام نختار القيمة 7 حيث يوجد على يمينه صف من الاعداد، وبتقاطع هذا الصف مع العمود أسفل درجة حرية البسط نجد أن $F_{(0.05,1,7)} = 5.59$.

وبالمثل نجد أن $F_{(0.025,3,5)} = 7.76$ ، $F_{(0.1,12,8)} = 2.50$

مثال2.4: إذا كان $F_{(\alpha,9,8)} = 4.35$ أوجد قيمة α.

الحل/

بالبحث في جدول (4) الخاص بالتوزيع وفي كل حالات α عن القيم الموجودة في تقاطعات أعمدة درجة الحرية 9 الخاصة بالبسط مع القيم الموجودة في صفوف درجة الحرية 8 الخاصة بالمقام نجد ما يلي:

α	0.1	0.05	0.025	0.01
$F_{(\alpha,9,8)}$	2.56	3.39	4.36	5.91

لاحظ أن القيمة 4.3572 هي أقرب قيمة من قيمة المتغير العشوائي المعطاة وهي 4.3452 لذلك نأخذ قيمة α تساوي 0.025

في بعض الأحيان تكون قيمة α هي احدى القيم { 0.9 ، 0.95 ، 0.975 ، 0.99 }

وهي ليس من ضمن القيم المدرجة في جدول (4)؛ لذلك وحتى نتمكن من استخدام جدول (4) في ايجاد قيم F نذكر بدون برهان النظرية التالية.

نظرية 2.7: بفرض أن $n_1, n_2 > 0$ فإن:

$$\mathbf{F(1-\alpha, n_1-1, n_2-1) = \frac{1}{F(\alpha, n_2-1, n_1-1)}}$$

مثال 2.5: أوجد قيمة $F(0.95,7,5)$

الحل/

لاحظ أن $\mathbf{1-\alpha} = 0.95$ وهي ليس من ضمن القيم المدرجة في جدول (4) لذلك بالتطبيق المباشر لنظرية 2.7 نجد أن:

$$F(0.95,7,5) = \frac{1}{F(0.05,5,7)} = \frac{1}{3.97} = 0.253$$

من استخدامات التوزيع الهامة ما يلي:

1- اختبار تجانس تبايني عينتين مستقلتين.
2- اختبار معنوية معامل الارتباط المتعدد.
3- اختبار معنوية نموذج الانحدار الخطي المتعدد.
4- اختبار تجانس عدة تقديرات مستقلة لتباين مجتمع طبيعي.
5- الاستخدام الواسع في اسلوب تحليل التباين.

سنتحدث الآن عن بعض توزيعات المعاينة الهامة و المتعلقة بالوسط الحسابي و الفرق بين وسطين، النسبة و الفرق بين نسبتين، والتباين و النسبة بين تباينين.

توزيعات المعاينة عند مجتمعات التوزيع الطبيعي

توزيع الوسط الحسابي للعينة The sampling distribution of means

لو أخذنا جميع العينات الممكنة التي حجمها $\boldsymbol{n}$ من مجتمع حجمه $\boldsymbol{N}$ وحسبنا أوساطها ؛ نجد أن هذه الأوساط قد تختلف، مما يعني أن سلوك هذه الاوساط هو سلوك متغير عشوائي يرمز له بالرمز $\overline{\boldsymbol{X}}$ ، وبالتالي هذا المتغير العشوائي له توزيع احتمالي مستمداً من توزيع المجتمع الذي سُحِبَت منه العينات يسمى توزيع الوسط الحسابي للعينة وسنذكر في هذا الفصل بعض الحالات المختلفة للتوزيع وكيفية حساب الاحتمالات لهذا المتغير العشوائي، وسنذكر بعض النظريات الهامة بدون برهان والمتعلقة بهذا التوزيع.

نظرية2.8: توقع المتغير العشوائي $\bar{X}$ والذي يرمز له بالرمز $\mu_{\bar{X}}$ هو:

$$E(\bar{X}) = \mu_{\bar{X}} = \mu$$

حيث μ هي وسط المجتمع.

نظرية2.9:إذا كان لدينا مجتمعاً غير محدود الحجم (غير منتهي) أو محدود الحجم(منتهي) والسحب منه بالإرجاع، فإن تباين توزيع الوسط الحسابي للعينة والذي يرمز له بالرمز $\sigma_{\bar{X}}^2$ معطى بالعلاقة:

$$\boldsymbol{E[(\bar{X} - \mu)^2] = \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}}$$

حيث σ^2 هو تباين المجتمع و n هو حجم العينة.

ملاحظات/

- يطلق على الانحراف المعياري لتوزيع المعاينة اسم الخطأ المعياري (**Standard Error**)

- إذا كان حجم المجتمع محدود وكان السحب منه بالإرجاع، فيعتبر المجتمع هنا غير منتهي.

نظرية 2.10: إذا كان حجم المجتمع محدود قيمته N وكان السحب بدون ارجاع وكان حجم العينة $n \le N$ فإن

$$\boldsymbol{\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)}$$

نظرية2.11: إذا كان لدينا مجتمعاً بوسط μ وتباين σ^2 فإنه بغض النظر عن توزيع المجتمع يكون:

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathbf{N(0,1)}$$

مثال2.6: اثبت أن $E(\bar{X}) = \mu$ وذلك عند سحب عينة عشوائية بحجم 2 مشاهدة من المجتمع {2،3,4} مرة مع الارجاع ومرة بدون ارجاع.

الحل/

أولاً/ السحب مع الارجاع

نلاحظ أن عدد العينات التي يمكن سحبها من المجتمع في هذه الحالة هو9 $N^n = 3^2 = 9$

و الجدول التالي يبين العينات و اوساطها و احتمالاتها و توقع المتغير العشوائي $\bar{X}$:

العينات الممكنة	$\bar{x}_i$	$p(\bar{x}_i)$	$\bar{x}_i p(\bar{x}_i)$
{2,2}	2	$1/9$	$4/18$
{2,3}	2.5	$1/9$	$5/18$
{2,4}	3	$1/9$	$6/18$
{3,2}	2.5	$1/9$	$5/18$
{3,3}	3	$1/9$	$6/18$
{3,4}	3.5	$1/9$	$7/18$
{4,2}	3	$1/9$	$6/18$
{4,3}	3.5	$1/9$	$7/18$
{4,4}	4	$1/9$	$8/18$
المجموع	27		3

وحيث أن:

$$E(\bar{X}) = \sum \bar{x}_i p(\bar{x}_i) = 3$$

$$\mu = \frac{\sum X}{N} = \frac{2+3+4}{3} = 3$$

إذن $E(\bar{X}) = \mu$

ثانياً/ السحب بدون ارجاع

نلاحظ أن عدد العينات التي يمكن سحبها من المجتمع في هذه الحالة هو

$$\binom{N}{n} = \binom{3}{2} = 3$$

و الجدول التالي يبين العينات و اوساطها و احتمالاتها و توقع المتغير العشوائي $\bar{X}$

العينات الممكنة	$\bar{x}_i$	$\mathrm{p}(\bar{x}_i)$	$\bar{x}_i \mathrm{p}(\bar{x}_i)$
{2,3}	2.5	$1/3$	$5/6$
{2,4}	3	$1/3$	$6/6$
{3,4}	3.5	$1/3$	$7/6$
المجموع	9		3

وحيث أن:

$$E(\bar{X}) = \sum \bar{x}_i \mathrm{p}(\bar{x}_i) = 3$$

$$\mu = \frac{\sum X}{N} = \frac{2+3+4}{3} = 3$$

إذن $\mathrm{E}(\bar{X}) = \mu$

مثال2.7: لمثال2.4 اثبت أن:

أولاً/ إذا كان السحب مع الارجاع فإن $\boldsymbol{\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}}$

ثانياً/ إذا كان السحب بدون ارجاع فإن $\boldsymbol{\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)}$

الحل/

نحسب أولاً تباين المجتمع:

$$\sigma^2 = \frac{\sum(X-\mu)^2}{N} = \frac{(2-3)^2+(3-3)^2+(4-3)^2}{3} = \frac{2}{3}$$

أولاً/ إذا كان السحب مع الارجاع

العينات الممكنة	$\bar{x}_i$	$p(\bar{x}_i)$	$\bar{x}_i p(\bar{x}_i)$	${\bar{x}_i}^2 p(\bar{x}_i)$
{2,2}	2	$1/9$	$4/18$	$16/36$
{2,3}	2.5	$1/9$	$5/18$	$25/36$
{2,4}	3	$1/9$	$6/18$	$36/36$
{3,2}	2.5	$1/9$	$5/18$	$25/36$
{3,3}	3	$1/9$	$6/18$	$36/36$
{3,4}	3.5	$1/9$	$7/18$	$49/36$
{4,2}	3	$1/9$	$6/18$	$36/36$
{4,3}	3.5	$1/9$	$7/18$	$49/36$
{4,4}	4	$1/9$	$8/18$	$64/36$
المجموع	27		3	$336/18$

$$\sigma^2_{\bar{X}} = E(\bar{X}^2) - \mu^2 = \frac{336}{18} - 3^2 = \frac{1}{3}$$

$$\frac{\sigma^2}{n} = \frac{2/3}{2} = \frac{1}{3}$$

إذن $\quad \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$

ثانياً/ إذا كان السحب بدون ارجاع

العينات الممكنة	$\bar{x}_i$	$p(\bar{x}_i)$	$\bar{x}_i p(\bar{x}_i)$	${\bar{x}_i}^2 p(\bar{x}_i)$
{2,3}	2.5	$1/3$	$5/6$	$25/12$
{2,4}	3	$1/3$	$6/6$	$36/12$
{3,4}	3.5	$1/3$	$7/6$	$49/12$
المجموع	27		3	$110/12$

$$\sigma_{\bar{X}}^2 = E(\bar{X}^2) - \mu^2 = \frac{110}{12} - 3^2 = \frac{1}{6}$$

$$\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) = \frac{2/3}{2}\left(\frac{3-2}{3-1}\right) = \frac{1}{6}$$

إذن $\quad \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$

سنتحدث الآن عن حالات مختلفة لإيجاد توزيع الوسط الحسابي للعينة.

أولاً/ إذا كان تباين المجتمع معلوم

إذا كان $X \sim N(\mu, \sigma^2)$ وسحبت منه عينة حجمها n فإن توزيع الوسط الحسابي للعينة هو:

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

ولحساب الاحتمالات للمتغير العشوائي $\bar{X}$ يتم إيجاد القيم المعيارية له حيث:

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$

مثال2.8: إذا كان $X \sim \mathrm{N}(9,4)$، فما هو التوزيع الاحتمالي لمتوسط عينة عشوائية بحجم 25 من مجتمع X. احسب $P(\bar{x} > 10)$.

الحل/

حيث أن $X \sim \mathrm{N}(9,4)$ إذن $\bar{X} \sim \mathrm{N}\left(9, \frac{4}{25}\right)$.

ولحساب الاحتمال المطلوب لاحظ أن $Z = \frac{\bar{X}-9}{2/\sqrt{25}} = \frac{\bar{X}-9}{0.4} \sim \mathrm{N}(0,1)$

وبالتالي يكون الاحتمال المطلوب:

$$P(\bar{x} > 10) = P\left(\frac{\bar{X} - 9}{2/\sqrt{25}} > \frac{10 - 9}{2/5}\right) = P(Z > 2.5)$$

$$= 1 - P(Z < 2.5) = 1 - 0.9938 = 0.0062$$

ثانياً/ إذا كان تباين المجتمع الطبيعي مجهول

في هذه الحالة نقدر تباين المجتمع بتباين العينة المسحوبة منه ونفرق هنا بين حالتين:

- **عندما يكون حجم العينة $n \geq 30$**

يكون توزيع الوسط الحسابي للعينة هو:

$$\bar{X} \sim \mathrm{N}\left(\mu, \frac{S^2}{n}\right)$$

حيث S^2 هو تباين العينة ويحسب من العلاقة $S^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$ ويكون التوزيع المعياري المقابل هو

$$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim N(0,1)$$

مثال2.9: إذا كان $X \sim \mathrm{N}(7, \sigma^2)$، فما التوزيع الاحتمالي لمتوسط عينة عشوائية بحجم 36 من مجتمع X تباينها 9، ثم احسب $P(\bar{x} > 8)$.

الحل/

حيث أن تباين المجتمع مجهول فإن: $\bar{X} \sim \mathrm{N}\left(\mu, \frac{S^2}{n}\right) = \mathrm{N}\left(7, \frac{9}{36}\right)$.

ولحساب الاحتمال المطلوب لاحظ أن: $Z = \frac{\bar{X} - 7}{3/\sqrt{36}} = \frac{\bar{X} - 7}{0.5} \sim \mathrm{N}(0,1)$

وبالتالي يكون الاحتمال المطلوب:

$$P(\bar{x} > 8) = P\left(Z > \frac{8-7}{0.5}\right) = 1 - P(Z < 2) = 1 - 0.9772 = 0.0228$$

- **عندما يكون حجم العينة $n < 30$**

يكون توزيع الوسط الحسابي للعينة هو:

(راجع نظرية 2.6)

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n-1}} \sim t(n-1)$$

ملاحظة/ خلال هذا الكتاب عندما نقول عينة حجمها كبير فإننا نقصد أن حجمها أكبر من 30 مشاهدة وعندما نقول عينة حجمها صغير فإننا نقصد أن حجمها أقل من 30 مالم نوضح في السياق غير ذلك.

مثال2.10: إذا كان $X \sim N(8.5, \sigma^2)$ أوجد التوزيع الاحتمالي لوسط العينة {9,10,9,8,8,7,8,5,8} ثم احسب $P(7 < \bar{x} < 9)$.

الحل/

بما أن تباين مجتمع X مجهول و حجم العينة $n = 9$ وهو أقل من 30 يكون توزيع الوسط الحسابي للعينة هو $\bar{X} \sim t(8)$.

ولحساب الاحتمال المطلوب نجد أولا تباين العينة:

$$\bar{x} = \frac{\sum x}{n} = \frac{72}{9} = 8$$

$$S^2 = \frac{\sum (x-8)^2}{8} = 2$$

وبالتالي يكون التوزيع:

$$T = \frac{\bar{X} - 8.5}{1.414/\sqrt{8}} \sim t(8)$$

ويكون الاحتمال المطلوب

$$P(7 < \bar{x} < 9) = P\left(\frac{7-8.5}{1.414/\sqrt{8}} < \frac{\bar{X}-8.5}{1.414/\sqrt{8}} < \frac{9-8.5}{1.414/\sqrt{8}}\right)$$

$$= P(-3 < T < 1)$$

$$= P(T < 1) - P(T < -3)$$

وحيث أن التوزيع متماثل حول الصفر، إذن $P(T < -3) = P(T > 3)$ ، و من قوانين الاحتمالات نعلم أن $P(T < 1) = 1 - P(T > 1)$ لذلك يكون المطلوب

$$P(7 < \bar{x} < 9) = 1 - P(T > 1) - P(T > 3)$$

$$= 1 - 0.1 - 0.01 = 0.89$$

ثالثاً/ إذا كان المجتمع طبيعي ذو حجم محدود و كان حجم العينة $n > 30$ وكان السحب بدون ارجاع فإنه:

- عند معلومية σ^2 يكون توزيع الوسط الحسابي للعينة هو:

$$\bar{\boldsymbol{X}} \sim \mathbf{N}\left(\boldsymbol{\mu}, \frac{\boldsymbol{\sigma}^2}{\boldsymbol{n}} \frac{\boldsymbol{N}-\boldsymbol{n}}{\boldsymbol{N}-\mathbf{1}}\right)$$

مثال2.11: من مجتمع بتوزيع طبيعي حجمه 200 مشاهدة وتباينه $\sigma^2 = 64$ تم سحب عينة عشوائية بدون إرجاع بحجم 50 مشاهدة، فما هو توزيع الوسط الحسابي للعينة، ثم احسب $p_r(\bar{x} > 32)$ عندما$\mu = 30$.

الحل/

حيث أن تباين المجتمع معلوم فإن توزيع الوسط الحسابي للعينة هو:

$$\bar{X}\sim \mathrm{N}\left(30,\frac{64}{50}\frac{200-50}{200-1}\right)=\mathrm{N}(30,0.96)$$

وبالتالي

$$Z=\frac{\bar{X}-30}{\sqrt{0.96}}\sim \mathrm{N}(0,1)$$

ويكون الاحتمال المطلوب هو:

$$P(\bar{x}>32)=P\left(Z>\frac{32-30}{\sqrt{0.96}}\right)=P(Z>2.02)$$

$$=1-P(Z<2.02)=1-0.9783=0.0217$$

- عند مجهولية σ^2 يكون توزيع الوسط الحسابي للعينة هو:

$$\bar{\boldsymbol{X}}\sim \mathbf{N}\left(\boldsymbol{\mu},\frac{\boldsymbol{S}^2}{\boldsymbol{n}}\frac{\boldsymbol{N}-\boldsymbol{n}}{\boldsymbol{N}-\mathbf{1}}\right)$$

حيث تم استبدال تباين العينة $\boldsymbol{S}^2$ بدلاً من تباين المجتمع $\boldsymbol{\sigma}^2$ في الحالة السابقة.

مثال2.12: أعد حل مثال2.11 إذا كان σ^2 مجهول وكان تباين العينة $S^2=49$.

الحل/

حيث أن تباين المجتمع مجهول فإن توزيع الوسط الحسابي للعينة هو:

$$\bar{X}\sim \mathrm{N}\left(30,\frac{49}{50}\frac{200-50}{200-1}\right)=\mathrm{N}(30,0.74)$$

وبالتالي

$$Z = \frac{\bar{X} - 30}{\sqrt{0.74}} \sim \mathrm{N}(0,1)$$

ويكون الاحتمال المطلوب هو:

$$P(\bar{x} > 32) = P\left(Z > \frac{32 - 30}{\sqrt{0.74}}\right) = P(Z > 2.32)$$

$$= 1 - P(Z < 2.32) = 1 - 0.9898 = 0.0102$$

ملاحظة/ عند سحب عينة من مجتمع توزيعه مجهول أو يتبع توزيعاً غير التوزيع الطبيعي، فإن توزيع الوسط الحسابي للعينة التي حجمها $n \geq 30$ يقرب إلى التوزيع الطبيعي وذلك حسب نظرية النهاية المركزية، حيث يكون:

$$\boldsymbol{\bar{X} \sim \mathrm{N}\left(\mu, \frac{S^2}{n}\right)}$$

مثال2.13: تم سحب عينة عشوائية بحجم 36 مشاهدة من مجتمع مجهول التوزيع، فبلغ الوسط الحسابي للعينة 10 بتباين 12 فما هو التوزيع الاحتمالي لمتوسط قيم العينة.

الحل/

مباشرة من الملاحظة السابقة يكون:

$$\bar{X} \sim \mathrm{N}\left(10, \frac{12}{36}\right)$$

توزيع الفرق بين متوسطي عينتين

إذا كان $X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$ وسحبت من مجتمعه عينة حجمها n_1 و $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$ وسحبت من مجتمعه عينة حجمها n_2 وكان المجتمعين مستقلين وعرفنا المتغيرين العشوائيين $\bar{X}_1$ و $\bar{X}_2$ فإن $\bar{X}_1 - \bar{X}_2$: يمثل متغيراً عشوائياً للفرق بين متوسطي العينتين. ومن نظرية 1.2 ونظرية 1.4 نحصل على الوسط الحسابي والتباين له وذلك كالتالي:

$$\boldsymbol{\mu_{\bar{X}_1-\bar{X}_2} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2}$$

$$\boldsymbol{\sigma^2_{\bar{X}_1-\bar{X}_2} = \sigma^2_{\bar{X}_1} + \sigma^2_{\bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

وسنتحدث الآن عن حالات مختلفة لإيجاد توزيع الفرق بين وسطي العينتين.

أولاً/ توزيع الفرق بين متوسطي عينتين مستقلتين.

وهنا نفرق بين ثلاث حالات:

- **عند معلومية σ_1^2 و σ_2^2**

$$\boldsymbol{\bar{X}_1 - \bar{X}_2 \sim \mathrm{N}\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}$$

وبالتالي:

$$\boldsymbol{Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim \mathrm{N}(0,1)}$$

مثال2.14: إذا كان $X_1 \sim \mathrm{N}(30,25)$ وسحبت من مجتمعه عينة حجمها 30 مشاهدة و $X_2 \sim \mathrm{N}(20,16)$ وسحبت من مجتمعه عينة حجمها 35 مشاهدة أوجد توزيع الفرق بين متوسطي العينتين. ثم احسب الاحتمال $P(\bar{x}_1 - \bar{x}_2 < 12)$.

الحل/

بالتطبيق المباشر للعلاقة

$$\bar{X}_1 - \bar{X}_2 \sim \mathrm{N}(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

نحصل على

$$\bar{X}_1 - \bar{X}_2 \sim \mathrm{N}\left(30 - 20, \frac{25}{30} + \frac{16}{35}\right) = \mathrm{N}(10,1.29)$$

وبالتالي فإن:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{1.29}} = \frac{(\bar{X}_1 - \bar{X}_2) - 10}{\sqrt{1.29}}$$

فيكون الاحتمال المطلوب هو

$$P(\bar{x}_1 - \bar{x}_2 < 12) = P\left(\frac{(\bar{X}_1 - \bar{X}_2) - 10}{\sqrt{1.29}} < \frac{12 - 10}{\sqrt{1.29}}\right)$$

$$= P\left(Z < \frac{12 - 10}{\sqrt{1.29}}\right)$$

$$= P(Z < 1.76) = 0.9608$$

- **عند مجهولية σ_1^2 و σ_2^2 وحجم العينتين كبير.**

$$\boldsymbol{\bar{X}_1 - \bar{X}_2 \sim \mathrm{N}(\mu_1 - \mu_2, \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2})}$$

مثال2.15: معمل ينتج 700 كغم من المكرونة كمعدل يومي، سحبت منه عينة عشوائية تمثل انتاج 40 يوماً فبلغ وسطها الحسابي 740 كغم بانحراف معياري 40 كغم . معمل آخر ينتج 500 كغم كمعدل يومي، سحبت منه عينة عشوائية تمثل انتاج 35 يوماً فبلغ وسطها الحسابي 480 كغم بانحراف معياري 20 كغم. فما هو التوزيع الاحتمالي للفرق بين متوسطي العينتين، ثم احسب الاحتمال التالي:

$$P(180 < \bar{X}_1 - \bar{X}_2 < 210)$$

الحل/

لدينا من المسألة $S_1^2 = 1600$ و $S_2^2 = 400$ و حيث أن $n_1 = 40$ ، $n_2 = 35$ وكليهما أكبر من 30 إذن يكون توزيع الفرق بين وسطي العينتين هو:

$$\bar{X}_1 - \bar{X}_2 \sim \mathrm{N}(700 - 500, \frac{1600}{40} + \frac{400}{35})$$

$$\bar{X}_1 - \bar{X}_2 \sim \mathrm{N}(200{,}51.43)$$

وبالتالي

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - 200}{\sqrt{51.43}} \sim \mathrm{N}(0{,}1)$$

لذلك يكون الاحتمال المطلوب هو

$$P(180 < \bar{X}_1 - \bar{X}_2 < 210) = P\left(\frac{180 - 200}{\sqrt{51.43}} < \frac{(\bar{X}_1 - \bar{X}_2) - 200}{\sqrt{51.43}} < \frac{210 - 200}{\sqrt{51.43}}\right)$$

$$= P(-2.79 < Z < 1.39) = P(Z < 1.39) - P(Z < -2.79)$$

$$= 0.9177 - 0.0029 = 0.9148$$

ملاحظة/ التوزيع أعلاه يصلح أيضاً عند السحب من مجتمعين غير طبيعيين أو مجهولي التوزيع بشرط كبر حجم العينة، وذلك استناداً لنظرية النهاية المركزية.

- **عند مجهولية تباين المجتمعين الطبيعيين وكان على الأقل حجم أحد العينتين صغير**

$$\bar{X}_1 - \bar{X}_2 \sim t(n_1 + n_2 - 2)$$

وسنميز هنا حالتين :

1- إذا كان $\sigma_1^2 = \sigma_2^2 = \sigma^2$

لاستنباط العلاقة التي سنستخدمها في هذه الحالة سنحسب اولا تباين توزيع الفرق بين المتوسطين في هذه الحالة كما يلي:

من نظرية النهاية المركزية نعلم أن

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \sim N(0,1)$$

من نظرية 2.4 يكون

$$\frac{(n_1 - 1)S_1^2}{\sigma^2} \sim \chi^2_{(n_1-1)} \quad \text{and} \quad \frac{(n_2 - 1)S_2^2}{\sigma^2} \sim \chi^2_{(n_2-1)}$$

من نظرية 2.2 يكون

$$\frac{(n_1 - 1)S_1^2}{\sigma^2} + \frac{(n_2 - 1)S_2^2}{\sigma^2} \sim \chi^2_{(n_1+n_2-2)}$$

بتطبيق نظرية 2.5 نحصل على

$$T = \frac{\dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}}{\sqrt{\left(\dfrac{(n_1 - 1)S_1^2}{\sigma^2} + \dfrac{(n_2 - 1)S_2^2}{\sigma^2}\right) \Big/ (n_1 + n_2 - 2)}} \sim t(n_1 + n_2 - 2)$$

بالتبسيط و بوضع

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

نحصل على

$$\boldsymbol{T} = \frac{(\bar{\boldsymbol{X}}_1 - \bar{\boldsymbol{X}}_2) - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)}{\sqrt{\boldsymbol{S}_p^2\left(\frac{\boldsymbol{1}}{\boldsymbol{n}_1} + \frac{\boldsymbol{1}}{\boldsymbol{n}_2}\right)}}$$

وبوضع

$$S_{\bar{X}_1-\bar{X}_2}^2 = S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right),$$

نحصل في النهاية على الصيغة المبسطة التالية:

$$\boldsymbol{T} = \frac{(\bar{\boldsymbol{X}}_1 - \bar{\boldsymbol{X}}_2) - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)}{\boldsymbol{S}_{\bar{X}_1-\bar{X}_2}} \sim \boldsymbol{t}(\boldsymbol{n}_1 + \boldsymbol{n}_2 - \boldsymbol{2})$$

مثال2.16: في مثال2.15 بفرض أن حجم العينة الاولى 20 يوماً بانحراف معياري 30 كغم وحجم العينة الثانية 18 يوما بانحراف معياري 25 كغم و بفرض تجانس تبايني المجتمعين أوجد $P(\bar{X}_1 - \bar{X}_2 > 215)$

الحل/

حيث أن $n_1 = 20$ ، $n_2 = 18$ وكليهما أصغر من 30 إذن توزيع الفرق بين وسطي العينتين هو:

$$\bar{X}_1 - \bar{X}_2 \sim \mathrm{t}(20 + 18 - 2)$$

$$\bar{X}_1 - \bar{X}_2 \sim \mathrm{t}(36)$$

ويكون تباين توزيع الفرق بين وسطي العينتين هو

$$S_p^2 = \frac{(20-1)30^2 + (18-1)25^2}{20+18-2} = 770.14$$

$$S_{\bar{X}_1-\bar{X}_2}^2 = 770.14\left(\frac{1}{20}+\frac{1}{18}\right) = 81.29$$

وبالتالي

$$\mathrm{T} = \frac{(\bar{X}_1-\bar{X}_2)-200}{\sqrt{81.29}} \sim t(36)$$

$$P(\bar{X}_1-\bar{X}_2 > 215) = P\left(\mathrm{T} > \frac{215-200}{\sqrt{81.29}}\right) = P(\mathrm{T} > 1.6637) = 0.05$$

2- إذا كان $\sigma_1^2 \neq \sigma_2^2$

$$\mathbf{T} = \frac{(\bar{\boldsymbol{X}}_1-\bar{\boldsymbol{X}}_2)-(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)}{\boldsymbol{S}_{\bar{X}_1-\bar{X}_2}} \sim \mathbf{t(f)}$$

حيث

$$\boldsymbol{S}_{\bar{X}_1-\bar{X}_2}^2 = \frac{\boldsymbol{S}_1^2}{\boldsymbol{n}_1}+\frac{\boldsymbol{S}_2^2}{\boldsymbol{n}_2}$$

وتحسب درجة الحرية من العلاقة:

$$\mathbf{f} = \frac{\left[\frac{\boldsymbol{S}_1^2}{\boldsymbol{n}_1}+\frac{\boldsymbol{S}_2^2}{\boldsymbol{n}_2}\right]^2}{\frac{\left(\boldsymbol{S}_1^2/\boldsymbol{n}_1\right)^2}{\boldsymbol{n}_1-\mathbf{1}}+\frac{\left(\boldsymbol{S}_2^2/\boldsymbol{n}_2\right)^2}{\boldsymbol{n}_2-\mathbf{1}}}$$

مثال2.17: إذا كان $X_1 \sim N(\mu_1, \sigma_1^2)$ وسحبت من مجتمعه عينة حجمها $n_1 = 40$ بتباين 9 و $X_2 \sim N(\mu_2, \sigma_2^2)$ وسحبت من مجتمعه عينة حجمها $n_2 = 15$ بتباين 16 أوجد توزيع الفرق بين وسطي العينتين، ثم احسب الاحتمال $P(\bar{X}_1 - \bar{X}_2 > 53)$ بفرض أن $\mu_1 = 300$ و $\mu_2 = 350$.

الحل/

حيث أن $n_1 = 40$ ، $n_2 = 15 < 30$ إذن توزيع الفرق بين وسطي العينتين هو:

$$f = \frac{\left[\frac{9}{40} + \frac{16}{15}\right]^2}{\frac{\left({}^{9}/_{40}\right)^2}{39} + \frac{\left({}^{16}/_{15}\right)^2}{14}} = \frac{(1.292)^2}{0.0013 + 0.0813} = 20.209$$

فنأخذ عدد درجات الحرية 20

$$\bar{X}_1 - \bar{X}_2 \sim t(20)$$

$$\bar{X}_1 - \bar{X}_2 \sim t(20)$$

ويكون تباين توزيع الفرق من وسطي العينتين هو

$$S^2_{\bar{X}_1 - \bar{X}_2} = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} = \frac{9}{40} + \frac{16}{15} = 1.29$$

وبناءً عليه فإن

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (350 - 300)}{1.14} \sim t(20)$$

ويكون الاحتمال المطلوب

$$P(\bar{X}_1-\bar{X}_2>53)=P\left(\mathrm{T}>\frac{53-50}{1.14}\right)=P(\mathrm{T}>2.63)=0.01$$

لاحظ هنا أن القيمة 2.63 غير موجودة في جدول (2) الخاص بتوزيع t لذلك نأخذ أقرب قيمة منها وهي 2.528

ملاحظة/ عند دراسة توزيع الفرق بين وسطي عينتين يفضل طرح الوسط الحسابي الاصغر من الاكبر حتى تكون النتيجة خالية من الاشارة ويبقى التعبير عن الفرق تعبيرا مطلقا.

ثانياً/ توزيع الفرق بين متوسطي عينتين مرتبطتين.

أحياناً تكون العينتان اللتان يتم سحبهما مرتبطتين، فمثلاً لقياس فاعلية برنامج تدريبي على عينة من الطلبة يتم قياس مستوى الطلبة قبل التعرض للبرنامج و بعده فتكون القراءات قبل وبعد التعرض للبرنامج تشكل عينتين مرتبطتين.

ولكي نوجد توزيع الفرق بين متوسطي المجتمعين الذين سحبنا منهما العينتين نلجأ لتحويل المسألة لإيجاد توزيع لمتوسط واحد وسنوضح كيف تتم هذه العملية.

بفرض أن $X\sim \mathrm{N}(\mu_1,\sigma_1^2)$ و $Y\sim \mathrm{N}(\mu_2,\sigma_2^2)$ وكان المجتمعان مرتبطان وسحبت من كل مجتمع عينة حجمها n $\{X_1,X_2,\ldots\ldots,X_n\}$ و $\{Y_1,Y_2,\ldots\ldots,Y_n\}$ على الترتيب، بحيث تمثل X_i و Y_i القيمتين المتناظرتين في العينتين، إذا كانت الفروق بين قيم العينتين المتناظرة هي $\{D_1,D_2,\ldots\ldots,D_n\}$ حيث $D_i=X_i-Y_i$ لجميع قيم $i=1,2,\ldots\ldots,n$ ، فإن $\{D_1,D_2,\ldots\ldots,D_n\}$ تشكل عينة الفروق ويمكن النظر لهذه العينة التي حجمها n على انها عينة عشوائية من مجتمع طبيعي وسطه $\mu_D=\mu_1-\mu_2$ وتباينه σ_D^2، ولنفرض أن وسط العينة μ_d و تباينها S_d^2. وسنميز هنا حالتين:

- **حجم العينة $n \geq 30$**

$$\bar{D} \sim N(\mu_D, \frac{\sigma_D^2}{n})$$

وبمعلومية تباين عينة الفروق ووسطها كمقدرين لوسط مجتمع الفروق وتباينه يكون:

$$Z = \frac{\bar{D} - \mu_d}{S_d / \sqrt{n}} \sim N(0,1)$$

- **حجم العينة $n < 30$**

$$\mathrm{T} = \frac{\bar{D} - \mu_d}{S_d / \sqrt{n-1}} \sim t(n-1)$$

وكمثال على هذا التوزيع قياس فعالية دواء معين لمعالجة مرض السكر على عينة معينة، حيث تقاس نسبة السكر قبل اخضاع المرضى للعلاج ثم بعد اخضاعهم للعلاج.

مثال2.18: في تجربة لبيان تحسن أداء العمال، تم سحب عينة عشوائية بحجم 16عاملا في المعمل فكانت قياس الكفاءة قبل وبعد دخولهم دورة تحسين الاداء كما هو مبين في الجدول التالي:

بعد (X_i)	7	8	7	7	8	8	9	9	9	8	7	7	7	8	8	8
قبل (Y_i)	5	5	5	4	7	8	5	6	7	8	5	6	5	7	7	5

اوجد توزيع الوسط الحسابي للعينة واحسب احتمال أن الفرق في الاداء قبل وبعد الدورة لا يقل عن 2.5

الحل/

نحسب الفروق $D_i = X_i - Y_i$ كما هو موضح في الجدول التالي ثم نجد المطلوب:

بعد (X_i)	7	8	7	7	8	8	9	9	9	8	7	7	7	8	8	8
قبل (Y_i)	5	5	5	4	7	8	5	6	7	8	5	6	5	7	7	5
D_i	2	3	2	3	1	0	4	3	2	0	2	1	2	1	1	3

$$\mu_d = \frac{\sum D_i}{n} = 1.875$$

$$S_d = \sqrt{\frac{\sum (D_i - \mu_d)^2}{n-1}} = 0.92$$

$$\mathrm{T} = \frac{\bar{D} - \mu_d}{S_d / \sqrt{n-1}} = \frac{\bar{D} - 1.875}{0.92 / \sqrt{15}} = \frac{\bar{D} - 1.875}{0.24} \sim t(15)$$

$$P(\bar{D} > 2.5) = P\left(\mathrm{T} > \frac{2.5 - 1.875}{0.24}\right) = P\left(\mathrm{T} > \frac{2.5 - 1.875}{0.24}\right)$$

$$= P(\mathrm{T} > 2.604) = 0.01$$

توزيع النسبة للعينة The sampling distribution of proportion

في المجتمعات الاحصائية التي تتبع توزيع ذي الحدين يتم رصد نسبة لخاصية معينة ندرسها في المجتمع الاحصائي ونرمز لها بالرمز **P**، مثل نسبة المدخنين ونسبة الصالح من انتاج مصنع معين،...الخ، حيث يتم حساب النسبة بقسمة مجموع مفردات الخاصية X على حجم المجتمع N أي أن $\mathbf{P}=\frac{X}{N}$ ، فلو أخذنا جميع العينات التي حجمها n من مجتمع حجمه N قد تختلف النسب لهذه العينات، مما يعني أن سلوك هذه النسب هو سلوك متغير عشوائي وبالتالي هذا المتغير العشوائي له توزيع احتمالي مستمدا من توزيع المجتمع الذي سحبت منه العينات، وسنرمز لنسبة الظاهرة في العينة بالرمز $\hat{\mathbf{p}}$ و بعدد مفردات الخاصية فيها x فتكون نسبة الخاصية في العينة $\hat{\mathbf{p}}=\frac{x}{n}$.

إن هذه النسبة تمثل متغير عشوائي له توزيع احتمالي يسمى توزيع النسبة حيث يكون وسطه الحسابي:

$$\mu_{\hat{\mathrm{p}}}=\mathrm{E}(\hat{\mathbf{p}})=\mathbf{E}\left(\frac{X}{n}\right)=\frac{1}{n}\mathbf{E}(X)=\mathbf{P}$$

ويكون تباينه هو:

$$\sigma_{\hat{\mathrm{p}}}^{2}=\mathrm{V}(\hat{\mathbf{p}})=\mathbf{V}\left(\frac{X}{n}\right)=\frac{1}{n^{2}}\mathbf{V}(X)=\frac{1}{n^{2}}\times n\mathbf{p}(1-\mathbf{p})=\frac{\mathbf{p}(1-\mathbf{p})}{n}$$

وفي حالة العينات الكبيرة $n\geq 30$ – وهي موضوع دراستنا – يقرب التوزيع للتوزيع الطبيعي وذلك استنادا لنظرية النهاية المركزية ونعبر عن ذلك بالصيغة

$$Z=\frac{\hat{\mathbf{p}}-\mu_{\hat{\mathbf{p}}}}{\sigma_{\hat{\mathbf{p}}}}\sim\mathbf{N}(0,1)$$

مثال2.19: مصنع ينتج عادة 25% عبوات كبيرة الحجم. سحبت من انتاجه عينة حجمها 2200 عبوة تبين أن منها 500 عبوة كبيرة الحجم. اوجد توزيع النسبة للعبوات الكبيرة ثم احسب احتمال أن المصنع ينتج أقل من 26% من العبوات الكبيرة في فترة اجراء البحث.

الحل/

لاحظ أن المجتمع هنا هو مجتمع ذي الحدين بنسبة نجاح $p = 0.25$ ، ولاحظ أن حجم العينة كبير وبحساب وسط التوزيع وتباينه نجد أن:

$$\mu_{\hat{p}} = P = 0.25$$

$$\sigma_{\hat{p}}^2 = \frac{p(1-p)}{n} = \frac{0.25 \times 0.75}{500} = 0.0004$$

$$\hat{p} \sim N(0.25, 0.0004)$$

$$Z = \frac{\hat{p} - \mu_{\hat{p}}}{\sigma_{\hat{p}}} = \frac{\hat{p} - 0.25}{0.02} \sim N(0,1)$$

$$P(\hat{p} < 0.26) = P\left(Z < \frac{0.26 - 0.25}{0.02}\right) = P(z < 0.5) = 0.6915$$

ملاحظة/ في حالة مجهولية p نعوض بدلا عنها بقيمة $\hat{p}$.

توزيع الفرق بين نسبتين

بفرض أن $X_1 \sim b(N_1, P_1)$ وسحبت من مجتمعه عينة حجمها $n_1 \geq 30$ و $X_2 \sim b(N_2, P_2)$ وسحبت من مجتمعه عينة حجمها $n_2 \geq 30$ وكان المجتمعين مستقلين فإن: $\hat{p}_1 - \hat{p}_2$ يمثل متغيراً عشوائياً للفرق بين نسبتي العينتين. ومن نظرية 1.2 ونظرية 1.4 نحصل على الوسط الحسابي والتباين له وذلك كالتالي:

$$\boldsymbol{\mu_{\hat{p}_1 - \hat{p}_2} = \mu_{\hat{p}_1} - \mu_{\hat{p}_2} = P_1 - P_2}$$

$$\boldsymbol{\sigma_{\hat{p}_1 - \hat{p}_2}^2 = \sigma_{\hat{p}_1}^2 + \sigma_{\hat{p}_2}^2 = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

ويكون توزيع فرق النسب

$$\mathbf{Z} = \frac{(\hat{\mathbf{p}}_1 - \hat{\mathbf{p}}_2) - (\mathbf{P}_1 - \mathbf{P}_2)}{\sqrt{\frac{\mathbf{p}_1(1 - \mathbf{p}_1)}{n_1} + \frac{\mathbf{p}_2(1 - \mathbf{p}_2)}{n_2}}} \sim \mathbf{N}(\mathbf{0}, \mathbf{1})$$

مثال2.20:إذا علم أن نسبة الذكور في مؤسسة A تبلغ 0.3 وفي مؤسسة B تبلغ 0.2، فإذا سحبت عينتين عشوائيا الاولى من المؤسسة A بحجم 100 و الثانية من المؤسسة B بحجم 200 فما احتمال أن يكون الفرق بين نسبتي الذكور في العينتين أكبر من 6%.

الحل/

لاحظ أن التوزيع هو توزيع فرق بين نسبتين بوسط حسابي وتباين:

$$\mu_{\hat{p}_1 - \hat{p}_2} = P_1 - P_2 = 0.3 - 0.2 = 0.1$$

$$\sigma^2_{\hat{p}_1 - \hat{p}_2} = \sigma^2_{\hat{p}_1} + \sigma^2_{\hat{p}_2} = \frac{0.3 \times 0.7}{100} + \frac{0.2 \times 0.8}{200} = 0.0029$$

ويكون توزيع فرق النسب

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - 0.1}{0.054} \sim N(0,1)$$

فيكون الاحتمال المطلوب هو

$$P(\hat{p}_1 - \hat{p}_2 > 0.06) = P\left(Z > \frac{0.06 - 0.1}{0.054}\right)$$

$$= P(Z > -0.74) = 1 - P(Z < -0.74)$$

$$= 1 - 0.2296 = 0.7704$$

توزيع التباين للعينة The sampling distribution of variance

من نظرية 2.4 يكون توزيع التباين هو

$$\mathbf{C} = \frac{(\boldsymbol{n}-\mathbf{1})\boldsymbol{S}^{\mathbf{2}}}{\boldsymbol{\sigma}^{\mathbf{2}}} \sim \chi^2_{(\boldsymbol{n}-\mathbf{1})}$$

وكتطبيق على هذا التوزيع سنعرض المثال التالي:

مثال2.21: سحبت عينة عشوائية بحجم 20 مشاهدة من مجتمع توزيع طبيعي تباينه 9، احسب احتمال أن يزيد تباين العينة عن 15

الحل/

حيث أن

$$\mathrm{C} = \frac{19S^2}{9} \sim \chi^2_{(19)}$$

إذن الاحتمال المطلوب

$$P(S^2 > 15) = P\left(\frac{19S^2}{9} > \frac{19 \times 15}{9}\right) = P(\mathrm{C} > 31.67) = 0.025$$

لحظ أن القيمة 31.67 غير موجودة في جدول (3) الخاص بالتوزيع فنأخذ أقرب قيمة لها وهي 32.852

توزيع النسبة بين تباين عينتين

يعتبر هذا التوزيع من التوزيعات الهامة التي تبحث في تجانس المجتمعات ونلجأ لحساب النسب بين التباينات وليس الفرق بينها لسهولة دراسة النسب و تفسيرها.

فإذا سحبت عينة حجمها n_1 وتباينها S_1^2 من مجتمع $\mathrm{N}(\mu_1, \sigma_1^2)$ وعينة أخرى حجمها n_2 وتباينها S_2^2 من مجتمع $\mathrm{N}(\mu_2, \sigma_2^2)$ مستقل عن المجتمع الاول فإن:

$$\frac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2_{(n_1-1)} \quad , \quad \frac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2_{(n_2-1)}$$

و بالتالي:

$$\frac{\frac{(n_1-1)S_1^2}{\sigma_1^2} \Big/ (n_1-1)}{\frac{(n_2-1)S_2^2}{\sigma_2^2} \Big/ (n_2-1)} \sim \mathrm{F}(n_1-1, n_2-1)$$

بعد تبسيط الطرف الأيسر نحصل على

$$\mathrm{F} = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim \mathrm{F}(n_1-1, n_2-1)$$

وإذا تساوى تبايني المجتمعين فإن

$$\frac{S_1^2}{S_2^2} \sim \mathrm{F}(n_1-1, n_2-1)$$

مثال2.22: سحبت عينة حجمها 13 من مجتمع طبيعي تباينه 9، وسحبت عينة أخرى حجمها 21 من مجتمع طبيعي تباينه 25 مستقل عن المجتمع الأول. أوجد احتمال النسبة بين تبايني العينتين أقل من 0.8

الحل/

حيث أن

$$\mathrm{F} = \frac{S_1^2/_9}{S_2^2/_{25}} \sim \mathrm{F}(12,20)$$

إذن

$$P\left(\frac{S_1^2}{S_2^2} < 0.8\right) = P\left(\frac{S_1^2/_9}{S_2^2/_{25}} < 0.8\left(\frac{25}{9}\right)\right)$$

$$= P(\mathrm{F} < 2.22)) = 1 - P(\mathrm{F} > 2.22))1 - 0.05 = 0.95$$

تمارين:

1- إذا كانت $X \sim t(18)$ احسب:

- $P(X > 1.338)$
- قيمة k عندما $P(X > k) = 0.05$

2- إذا كانت $X \sim \chi^2_{(5)}$ احسب:

- $P(X > 12.38)$
- $P(X < 9.24)$
- قيمة k عندما $P(X < 2k) = 0.025$

3- إذا كان $X_1 \sim \chi^2_{(10)}$ ، $X_2 \sim \chi^2_{(100)}$ اجب عن ما يلي:

- احسب $P(X_1 > 9)$ ، $P(X_2 < 85)$
- هل يمكن استخدام جدول **z** لإيجاد $P(X_2 < 85)$ علل اجابتك.

4- إذا كانت أطوال شجيرات المزرعة **A** تتبع توزيع طبيعي بوسط حسابي 70سم و تباين σ^2 . فإذا سحبت عينة عشوائية مكونة من 20 شجيرة، فما احتمال أن يكون متوسط طول الشجيرة في العينة أقل من 68سم، إذا علمت أن الانحراف المعياري للعينة 4سم.

5- أعد حل سؤال 4 بفرض أن تباين المجتمع يساوي 25

6- عينة عشوائية بحجم 25 مشاهدة مسحوبة من مجتمع طبيعي التوزيع بوسط حسابي **μ** وتباين 16 فما قيمة الاحتمالات التالية

- $P(S^2 > 12.2)$
- $P(S^2 < 0.67)$
- $P(11.2 < S^2 < 20.5)$

7- إذا علمت أن عمر جهاز كهربائي معين يتبع توزيع طبيعي بوسط 1000 ساعة و تباين 900، سحبت عينة عشوائية حجمها 50 جهازا ، احسب احتمال ان متوسط عمر الجهاز في العينة أقل من متوسط المجتمع بنسبة 15%.

الفصل الثالث

نظرية التقدير

Estimation theory

تلعب نظرية التقدير دورً رئيساً هاماً في الاحصاء الاستنتاجي، حيث يتم على ضوئها تقدير معالم المجتمع الاحصائي (Parameters) والمعلوم توزيعه الاحتمالي وذلك عن طريق سحب عينة عشوائية من المجتمع وتستخدم إحصاءاتها (Statistics) في تقدير معالم المجتمع.

فإذا كان لدينا متغير عشوائي X له اقتران احتمالي $f(x,\theta)$ حيث θ تمثل معلمة توزيع ذلك المتغير، ونريد تقدير هذه المعلمة، فنقوم بإيجاد مقدر **Estimator** لتقدير هذه المعلمة من خلال مشاهدات عينة عشوائية نسحبها من قيم المتغير العشوائي، فيكون المقدر نفسه متغير عشوائي حيث تتغير قيمته من عينة لأخرى ويرمز لهذا المقدر بالرمز $\hat{\theta}$.

وتتم عملية التقدير بطريقتين:

1- التقدير النقطي (**Point Estimation**) ويقصد به تمثيل معلمة المجتمع بقيمة ما نحسبها لعينة عشوائية باستخدام المقدر، فمثلا نقدر العمر الافتراضي للمصابيح المنتجة في مصنع معين بأن نقول أن العمر الافتراضي للمصباح يقدر بـ 300 ساعة. وهناك طرق عديدة لإيجاد مقدرات نقطية لمعالم المجتمع نذكر منها دون التطرق لبراهين على سبيل المثال:

- طريقة العزوم **Moment Method**
- طريقة الجوازية العظمى **Maximum Likelihood Method**

فإذا أردنا تقدير وسط مجتمع يتبع التوزيع الطبيعي بوسط عينة $\{\boldsymbol{X_i}\}_{\boldsymbol{i=1}}^{\boldsymbol{n}}$ سُحِبت منه، نستخدم إحدى الطريقتين للحصول على المقدر، حيث باستخدام الطريقتين السابقتين نحصل على المقدر التالي:

$$\overline{\boldsymbol{X}} = \frac{\sum \boldsymbol{X_i}}{\boldsymbol{n}}$$

2- التقدير بفترة (**Interval Estimation**) حيث تقدر معلمة التوزيع بفترة نطلق عليها اسم فترة الثقة **Confidence Intervals.**

فمثلاً ممكن أن نقدر العمر الافتراضي للمصابيح المنتجة في مصنع معين بأن نقول أن العمر الافتراضي للمصباح بالساعة يقع في الفترة [300,350].

خواص المقدر الجيد

حتى نقول أن المقدر هو مقدر جيد يجب أن يحقق الشروط التالية:

1− عدم التحيز Un-biasedness

يعتبر المقدر غير متحيز إذا تحقق الشرط: $\mathbf{E(\widehat{\theta}) = \theta}$

مثال3.1:المقدر $\bar{X} = \frac{\sum X_i}{n}$ هو مقدر غير متحيز لوسط مجتمع يتبع التوزيع الطبيعي **μ** حيث:

$$\mathrm{E}(\bar{X}) = E\left(\frac{\sum X_i}{n}\right) = \frac{1}{n}\left(\sum E(X_i)\right) = \frac{n\mu}{n} = \mu$$

مثال3.2: يعتبر المقدر التالي:

$$\hat{S}^2 = \frac{\sum (X_i - \bar{X})^2}{n}$$

مقدرا متحيزا لتباين المجتمع الطبيعي الذي تباينه σ^2 وذلك لأن:

$$\begin{aligned}
\mathrm{E}(\hat{S}^2) &= E\left(\frac{\sum (X_i - \bar{X})^2}{n}\right) = \frac{1}{n} E\left(\sum [(X_i - \mu) - (\bar{X} - \mu)]^2\right) \\
&= \frac{1}{n} E\left(\sum (X_i - \mu)^2 - 2(\bar{X} - \mu)\sum (X_i - \mu) + \sum (\bar{X} - \mu)^2\right) \\
&= \frac{1}{n} E\left(\sum (X_i - \mu)^2 - n(\bar{X} - \mu)^2\right) \\
&= \frac{1}{n}\left(\sum E(X_i - \mu)^2 - nE(\bar{X} - \mu)^2\right) \\
&= \frac{1}{n}\left(n\sigma^2 - n\frac{\sigma^2}{n}\right) = \frac{n-1}{n}\sigma^2
\end{aligned}$$

أما المقدر

$$S^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$$

فيعتبر مقدرا غير متحيزا لتباين المجتمع المذكور حيث يمكن بنفس الطريقة السابقة إثبات أن:

$$\mathrm{E}(S^2) = \sigma^2$$

2- الاتساق Consistency

نقول أن المقدر $\hat{\boldsymbol{\theta}}$ أنه متسق إذا اقتربت قيمة المقدر من قيمة المعلمة كلما كبر حجم العينة، وبصيغة أخرى:

$$\lim_{n\to\infty} p_r(|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}|) > \epsilon) = 0\,, \quad \epsilon > 0$$

والنظرية التالية تذكر صورة مكافئة للتعريف السابق للاتساق وي مفيدة في اثبات الاتساق.

نظرية 3.1: إذا كان $\lim\limits_{n\to\infty} \mathrm{E}(\hat{\theta}) = \theta$ و $\lim\limits_{n\to\infty} \mathrm{V}(\hat{\theta}) = 0$ فإن $\hat{\theta}$ مقدر متسق للمعلمة θ .

مثال3.3: لمجتمع طبيعي التوزيع بوسط μ وتباين σ^2 يعتبر المقدر $\bar{X} = \frac{\sum x_i}{n}$ مقدراً متسقاً للوسط الحسابي وذلك لأن:

$$\lim_{n\to\infty} \mathrm{E}(\bar{X}) = \lim_{n\to\infty} \mu = \mu$$

$$\lim_{n\to\infty} \mathrm{V}(\bar{X}) = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0$$

3- الكفاءة Efficiency

إذا كان لدينا أكثر من مقدر لمعلمة معينة فنقول أن المقدر الأكفأ هو المقدر الذي له أقل تباين، فالكفاءة خاصية نسبية تستخدم للمقارنة بين المقدرات المختلفة لنفس المعلمة وبنفس حجم العينة، ويكون المقدر الأكفأ هو الذي يعطي تقديرات قريبة من المعلمة المجهولة.

فإذا كان لدينا مقدرين $\hat{\theta}_1$ و $\hat{\theta}_2$ للمعلمتين θ_1 و θ_2 على الترتيب، فإن $\hat{\theta}_1$ أكفأ من $\hat{\theta}_2$ إذا كان $V(\hat{\theta}_1) < V(\hat{\theta}_2)$

ويتم حساب كفاءة المقدر **d** من العلاقة:

$$\boldsymbol{e}(\mathbf{d}) = \frac{\mathbf{1}}{\boldsymbol{V}(\mathbf{d})\mathbf{nE}\left(\frac{\boldsymbol{\partial} \ln \boldsymbol{f}(\boldsymbol{x},\boldsymbol{\theta})}{\boldsymbol{\partial\theta}}\right)^{\mathbf{2}}}$$

أو بصورة مكافئة

$$\boldsymbol{e}(\mathbf{d}) = \frac{\mathbf{1}}{\boldsymbol{V}(\mathbf{d})\mathbf{nE}\left(\frac{\boldsymbol{\partial}^{\mathbf{2}} \ln \boldsymbol{f}(\boldsymbol{x},\boldsymbol{\theta})}{\boldsymbol{\partial\theta}^{\mathbf{2}}}\right)}$$

مثال3.4: إذا كان X متغير عشوائي بحيث $X\sim \mathrm{b}(1,\theta)$ فابحث كفاءة المقدر $\mathrm{d}=\bar{X}=\frac{\sum X_i}{n}$ كمقدر للمعلمة θ.

الحل/

إن التوزيع المذكور هو توزيع برنولي وهو حالة خاصة من توزيع ذي الحدين عندما تكون $n=1$ حيث اقترانه الاحتمالي هو $f(x,\theta)=\theta^x(1-\theta)^{1-x}$ وتوقعه $\mathrm{E}(X)=\theta$ وتباينه $V(X)=\theta(1-\theta)$ و بالتالي فإن:

$$\ln f(x,\theta)=x\ln\theta+(1-x)\ln(1-\theta)$$

$$\frac{\partial \ln f(x,\theta)}{\partial\theta}=\frac{x}{\theta}-\frac{1-x}{1-\theta}=\frac{x-\theta}{\theta(1-\theta)}$$

$$\mathrm{E}\left(\frac{\partial \ln f(x,\theta)}{\partial\theta}\right)^2=\frac{1}{\theta^2(1-\theta)^2}\mathrm{E}(x-\theta)^2$$

$$=\frac{1}{\theta^2(1-\theta)^2}\mathrm{V(X)}=\frac{1}{\theta(1-\theta)}$$

$$V(\mathrm{d})=\frac{\theta(1-\theta)}{n}$$

$$e(\mathrm{d})=\frac{1}{\frac{\theta(1-\theta)}{n}.n.\frac{1}{\theta(1-\theta)}}=1$$

وتعتبر هذه القيمة جيدة وبالتالي يكون هذا المقدر هو الغير متحيز و الأكفأ.

4- **الكفاية Sufficiency**

يطلق على مقدر صفة الكفاية إذا كان يحوي جميع أو أغلب مشاهدات العينة، بحيث لا يوجد مقدر غيره يمكن أن يحوي تلك المشاهدات.

مثال3.5: الوسط الحسابي للعينة المسحوبة من مجتمع طبيعي هو مقدر كافٍ مقارنة بالوسيط والمنوال حيث أنه يعتمد على جميع المشاهدات بخلاف الوسيط والمنوال.

التقدير بفترة

أولاً/ تقدير وسط مجتمع بفترة (إيجاد فترة ثقة لوسط مجتمع)

في هذه الحالة لا نكتفي بتقدير معلمة المجتمع بنقطة فنلجأ باستخدام عينة من المجتمع لحساب فترة نأمل أن تقع فيها معلمة المجتمع بمستوى ثقة محدد مسبقا يرمز له بالرمز $(\boldsymbol{1-\alpha})$ حيث تسمى $\boldsymbol{\alpha}$ بمستوى المعنوية ويطلق على هذه الفترة فترة ثقة للمعلمة بمستوى ثقة $\boldsymbol{(1-\alpha)\%}$.

ملاحظة: خلال هذا الكتاب سنرمز لـ $\boldsymbol{P(Z>\alpha)}$ بالرمز $\boldsymbol{Z_\alpha}$

وفيما يلي سنوضح كيفية إيجاد فترة ثقة لوسط مجتمع طبيعي في حالات مختلفة:

1- **عند معلومية تباين المجتمع** σ^2.

نعلم أنه إذا كان $X\sim N(\mu,\sigma^2)$ فإن $\bar{X}\sim N(\mu,\frac{\sigma^2}{n})$ وكذلك $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$ ، فنعرف فترة ثقة $(1-\alpha)\%$ أو نقول **{ فترة ثقة عند مستوى معنوية** α **}** بأنها الفترة التي تحقق ما يلي:

$$\boldsymbol{P\left(-Z_{\frac{\alpha}{2}}<\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}<Z_{\frac{\alpha}{2}}\right)=1-\alpha}$$

وبتبسيط فترة الاحتمال نحصل على:

$$\boldsymbol{-Z_{\frac{\alpha}{2}}<\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}<Z_{\frac{\alpha}{2}}}$$

$$\boldsymbol{\bar{X}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}<\mu<\bar{X}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}}$$

حيث يسمى المقدار $\boldsymbol{Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}}$ حد $\boldsymbol{(1-\alpha)\%}$ للخطأ في تقدير وسط المجتمع وسنرمز له بالرمز $\mathbf{d}$ أي:

$$\mathbf{d}=\boldsymbol{Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}}$$

وهذه هي فترة الثقة المطلوبة، ولاحظ أن الفترة مركزها $\boldsymbol{\bar{X}}$ وحيث أن $\boldsymbol{\bar{X}}$ متغير عشوائي سيتغير مركز الفترة بتغير العينة وبالتالي ستتغير فترة الثقة، ولكن كل الفترات تشترك في أن وسط العينة $\boldsymbol{\mu}$ يقع داخل كل فترة بنسبة $(1-\alpha)\%$.

مثال3.5: أخذت عينة حجمها 49 ووسطها 45 من مجتمع إحصائي يتخذ توزيعاً طبيعياً تباينه 12.25 ، جد فترة ثقة 95% لوسط المجتمع.

الحل/

حجم العينة $\mathbf{n = 49}$ وحيث أن $\mathbf{1 - \propto = 0.95}$ فتكون $\mathbf{\propto = 0.05}$ ومنها $\frac{\propto}{2} = \mathbf{0.025}$ وبالتالي $\mathbf{Z_{\frac{\propto}{2}} = 1.96}$ وذلك من جدول المساحات للتوزيع الطبيعي المعياري. و حيث أن $\mathbf{\sigma^2 = 12.25}$ فيكون الانحراف المعياري $\mathbf{\sigma = 3.5}$ ، ولاحظ هنا أن تباين المجتمع معلوم لذلك سنستخدم التوزيع الطبيعي في ايجاد فترة الثقة المطلوبة عند مستوى دلالة 95% وذلك كما يلي:

$$\bar{x} - Z_{\frac{\propto}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\frac{\propto}{2}} \frac{\sigma}{\sqrt{n}}$$

$$45 - 1.96 \frac{3.5}{\sqrt{49}} \leq \mu \leq 45 + 1.96 \frac{3.5}{\sqrt{49}}$$

$$45 - 0.98 \leq \mu \leq 45 + 0.98$$

$$44.02 \leq \mu \leq 45.98$$

وهذا يعني أن احتمال وقوع وسط المجتمع بين القيمتين 45.95 ، 44.02 هو 95%. وتعني النسبة 95% أيضاً أن 5 حالات من كل 100 حالة ستكون فيها $\boldsymbol{\mu}$ خارج الفترة المذكورة.

مثال3.6: إذا كانت أجور العمال في إحدى المؤسسات تتبع في توزيعها التوزيع الطبيعي بانحراف معياري 20 دينار، أوجد فترة ثقة 95% باستخدام عينة الاجور {250،150،250،200،150،200،180،150،180،250}

الحل/

حجم العينة $n = 10$ وحيث أن $1 - \propto = 0.95$ إذن $\propto = 0.05$ ومنها $\frac{\propto}{2} = 0.025$ وبالتالي $Z_{\frac{\propto}{2}} = 1.96$ وذلك من جدول المساحات للتوزيع الطبيعي المعياري.

نحسب الوسط الحسابي للعينة

$$\bar{X} = \frac{\sum X_i}{n} = \frac{1960}{10} = 196$$

فتكون فترة الثقة المطلوبة

$$\bar{x} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$196 - 1.96 \frac{20}{\sqrt{10}} \leq \mu \leq 196 + 1.96 \frac{20}{\sqrt{10}}$$

$$183.6 \leq \mu \leq 208.4$$

وعليه يكون الحد الأدنى للأجور 183.6 و الحد الاقصى للأجور 208.4 وذلك عند مستوى الثقة المذكور.

2- عند مجهولية تباين المجتمع σ^2.

- إذا كان $n \geq 30$ فإن فترة الثقة لوسط المجتمع هي:

$$\bar{X} - Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} < \mu < \bar{X} + Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

حيث سنستخدم تباين العينة بدلا من تباين المجتمع.

مثال3.7:سحبت عينة عشوائية من مجتمع طبيعي مجهول التباين بحجم 36 مشاهدة حيث بلغ وسطها الحسابي 12 بتباين 16 ، اوجد فترة ثقة 99% لتقدير وسط المجتمع.

الحل/ من المسألة نجد أن $S^2 = 16 , \bar{x} = 12$ وحجم العينة $n = 36$ ، وحيث أن $1-\alpha = 0.99$ إذن $\alpha = 0.01$ ومنها $\frac{\alpha}{2} = 0.005$ وبالتالي $Z_{\frac{\alpha}{2}} = 2.58$ وذلك من جدول (1) للتوزيع الطبيعي المعياري.

فتكون فترة الثقة المطلوبة هي:

$$\bar{x} - Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} < \mu < \bar{x} + Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

$$12 - 2.58 \frac{4}{\sqrt{36}} < \mu < 12 + 2.58 \frac{4}{\sqrt{36}}$$

$$10.28 < \mu < 13.72$$

- إذا كان $\boldsymbol{n} < \mathbf{30}$ فإن فترة الثقة لوسط المجتمع هي:

$$\bar{X} - t_{\left(\frac{\alpha}{2}, n-1\right)} \frac{S}{\sqrt{n-1}} < \mu < \bar{X} + t_{\left(\frac{\alpha}{2}, n-1\right)} \frac{S}{\sqrt{n-1}}$$

مثال3.8: العينة $X: 10,9,11,6,8,7,10,8,9,7$ سحبت من مجتمع طبيعي مجهول التباين، أوجد فترة ثقة 95% لوسط هذا المجتمع.

الحل/

بحساب الوسط و الانحراف المعياري للعينة نجد أن $\bar{x} = 8.5$, $S = 1.58$ وحيث أن حجم العينة $n = 10$ أقل من 30 فإن التوزيع المستخدم هو توزيع t ومن جدول التوزيع نجد أن $t_{(0.025,9)} = 2.26$

وتكون فترة الثقة المطلوبة هي:

$$\bar{x} - t_{\left(\frac{\alpha}{2}, n-1\right)} \frac{S}{\sqrt{n-1}} < \mu < \bar{x} + t_{\left(\frac{\alpha}{2}, n-1\right)} \frac{S}{\sqrt{n-1}}$$

$$8.5 - 2.26 \frac{1.58}{\sqrt{9}} < \mu < 8.5 + 2.26 \frac{1.58}{\sqrt{9}}$$

$$7.31 < \mu < 9.69$$

تحديد الحجم اللازم للعينة لتقدير وسط المجتمع

عرفنا فيما سبق المقدار $\mathbf{d} = Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ على أنه حد $(1-\alpha)\%$ للخطأ في تقدير وسط المجتمع، و بفرض أنه حد الخطأ الاكبر المسموح به والذي يتم تحديده مسبقاً من قبل الباحث، فبحل المتباينة $Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mathbf{d}$ بالنسبة لحجم العينة n نجد حجم العينة المناسب لتحقيق هذا الحد وذلك كما يلي:

$$Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mathbf{d} \quad \Rightarrow \quad \frac{1}{\sqrt{n}} \leq \frac{\mathbf{d}}{\sigma Z_{\frac{\alpha}{2}}} \quad \Rightarrow \quad \sqrt{n} \geq \frac{\sigma Z_{\frac{\alpha}{2}}}{\mathbf{d}}$$

$$n \geq \left(\frac{\sigma Z_{\frac{\alpha}{2}}}{\mathbf{d}}\right)^2$$

وإذا كانت $\boldsymbol{\sigma}$ مجهولة نقدرها بوسط العينة.

مثال3.9: لاحظ مدرس بخبرته أن وسط درجات الطلاب في مادة الاحصاء 75 علامة وبانحراف معياري 9 علامات. إذا رغب المدرس في تطوير اسلوب تدريس المادة ومن ثم تقدير الوسط الحسابي للعلامات وفق الاسلوب الجديد بحيث يكون متأكداً بنسبة 95% أن الخطأ في المقدر الناتج لا يزيد عن 3 علامات، فكم طالباً يحتاج لإخضاعهم للتجربة.

الحل/ لاحظ أن $Z_{\frac{\alpha}{2}} = 1.96$ وبالتعويض المباشر في المتباينة السابقة نحصل على:

$$n \geq \left(\frac{9 \times 1.96}{3}\right)^2 = 34.57$$

فنأخذ $n = 35$ طالباً كحجم للعينة حتى نحصل على لا يزيد الخطأ في المقدر عن 3.

ثانياً/ تقدير الفرق بين وسطي مجتمعين مستقلين بفترة

وفيما يلي سنجد فترة الثقة للفرق بين وسطي مجتمعين طبيعيين مستقلين في حالات مختلفة:

1- **عند معلومية تباين المجتمعين** σ_1^2 و σ_2^2 .

إذا كان $X_1 \sim N(\mu_1, \sigma_1^2)$ و $X_2 \sim N(\mu_2, \sigma_2^2)$ فإن:

$$\bar{X}_1 - \bar{X}_2 \sim \mathrm{N}(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

ومن نظرية 2.1 نجد أن:

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

وتكون فترة ثقة $(1 - \alpha)\%$ كالتالي:

$$-\boldsymbol{Z_{\frac{\alpha}{2}}} < \frac{(\bar{\boldsymbol{X}}_1 - \bar{\boldsymbol{X}}_2) - (\boldsymbol{\mu_1} - \boldsymbol{\mu_2})}{\sqrt{\frac{\boldsymbol{\sigma_1^2}}{\boldsymbol{n_1}} + \frac{\boldsymbol{\sigma_2^2}}{\boldsymbol{n_2}}}} < \boldsymbol{Z_{\frac{\alpha}{2}}}$$

حيث بحل المتباينة بالنسبة إلى $\boldsymbol{\mu_1 - \mu_2}$ نجد أن:

$$(\bar{\boldsymbol{X}}_1 - \bar{\boldsymbol{X}}_2) - \boldsymbol{Z_{\frac{\alpha}{2}}}\sqrt{\frac{\boldsymbol{\sigma_1^2}}{\boldsymbol{n_1}} + \frac{\boldsymbol{\sigma_2^2}}{\boldsymbol{n_2}}} < \boldsymbol{\mu_1} - \boldsymbol{\mu_2} < (\bar{\boldsymbol{X}}_1 - \bar{\boldsymbol{X}}_2) + \boldsymbol{Z_{\frac{\alpha}{2}}}\sqrt{\frac{\boldsymbol{\sigma_1^2}}{\boldsymbol{n_1}} + \frac{\boldsymbol{\sigma_2^2}}{\boldsymbol{n_2}}}$$

مثال3.10: سُحبت عينتان من مجتمعين طبيعيين مستقلين وكانت كل عينة تمثل سرعة إنجاز تمرين معين و العينتان هما:

$$X_1: 20,18,22,22,18,20,25,25$$
$$X_2: 25,22,24,23,25,30,30,27,20$$

أوجد فترة ثقة 95% للفرق بين إنجازي المجموعتين في حالة تباين المجتمع الأول 6 و تباين المجتمع الثاني 12

الحل/

نحسب في البداية وسطي العينتين

$$\bar{X}_1 = \frac{\sum X_i}{n} = \frac{170}{8} = 21.25$$
$$\bar{X}_2 = \frac{\sum X_i}{n} = \frac{226}{9} = 25.11$$

معلوم لدينا $\sigma_1^2 = 6$ و $\sigma_2^2 = 12$ و حجم العينات $n_1 = 8$ و$n_2 = 9$ وحيث أن $1-\propto = 0.95$ إذن $\propto = 0.05$ ومنها $\frac{\propto}{2} = 0.025$ وبالتالي $Z_{\frac{\propto}{2}} = 1.96$ وذلك من جدول (1) للتوزيع الطبيعي المعياري.

فتكون فترة الثقة كما يلي:

$$(\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$
$$-3.86 - 1.96\sqrt{\frac{6}{8} + \frac{12}{9}} < \mu_1 - \mu_2 < -3.86 + 1.96\sqrt{\frac{6}{8} + \frac{12}{9}}$$
$$-6.69 < \mu_1 - \mu_2 < -1.03$$

2- **عند مجهولية تباين المجتمعين** σ_1^2 و σ_2^2 .

- إذا كان $\boldsymbol{n_1, n_2 \geq 30}$

في هذه الحالة نقدر تباين كل مجتمع بتباين عينة عشوائية تسحب منه ونستبدل تباين المجتمعين σ_1^2 , σ_2^2 بتباين العينتين S_1^2 , S_2^2 في فترة الثقة للحالة السابقة فتصبح فترة الثقة لهذه الحالة كما يلي:

$$(\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

مثال3.11: أوجد فترة ثقة 98% للفرق بين وسطي مجتمعين طبيعيين مستقلين إذا علمت أن: حجم العينة الاولى 160 ووسطها 81.2 والانحراف المعياري لها 7.6 و حجم العينة الثانية 90 ووسطها 76.4 و الانحراف المعياري لها 8.2

الحل/

من المسألة نجد أن :

$$n_1 = 160 \quad , \quad \bar{X}_1 = 81.2 \quad , \quad S_1^2 = 7.6$$
$$n_2 = 90 \quad , \quad \bar{X}_2 = 76.4 \quad , \quad S_2^2 = 8.2$$

ولاحظ أن $\alpha = 0.02$ ومن جدول (1) للتوزيع الطبيعي نجد أن $Z_{\frac{\alpha}{2}} = 2.33$ فتكون فترة الثقة المطلوبة كما يلي:

$$(\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$
$$4.8 - 2.33\sqrt{\frac{7.6}{160} + \frac{8.2}{90}} < \mu_1 - \mu_2 < 4.8 + 2.33\sqrt{\frac{7.6}{160} + \frac{8.2}{90}}$$
$$3.933 < \mu_1 - \mu_2 < 5.567$$

- إذا كان $n_1, n_2 < 30$ (أو أحدهما أقل من 30) وكان تباينا المجتمعين غير متساوٍ فإن فترة الثقة:

$$(\bar{X}_1 - \bar{X}_2) - t_{(\frac{\alpha}{2}, n-1)}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{(\frac{\alpha}{2}, n-1)}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

مثال3.12: أعد حل مثال3.9 في حالة مجهولية تباين المجتمعين.

معلوم لدينا من المثال المذكور أن $\bar{X}_1 = 21.25$ و $\bar{X}_2 = 25.11$ وحجوم العينات هو $n_1 = 8$ و$n_2 = 9$ وهما أقل من 30 و من جدول توزيع t نجد أن $t_{(0.025,15)} = 2.131$.

بحساب تباين العينتين نجد أن $S_1^2 = 7.618$ و $S_2^2 = 11.628$ وبالتالي تكون فترة الثقة المطلوبة هي:

$$(\bar{X}_1 - \bar{X}_2) - t_{(\frac{\alpha}{2},n-1)}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{(\frac{\alpha}{2},n-1)}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$-3.86 - 2.131\sqrt{\frac{7.618}{8} + \frac{11.628}{9}} < \mu_1 - \mu_2 < -3.86 + 2.131\sqrt{\frac{7.618}{8} + \frac{11.628}{9}}$$

$$-7.05 < \mu_1 - \mu_2 < -0.67$$

ثالثاً/ تقدير الفرق بين وسطي مجتمعين مرتبطين بفترة

وفيما يلي سنجد فترة الثقة للفرق بين وسطي مجتمعين طبيعيين مرتبطين.

1- في حالة $n \geq 30$ يكون

$$\boldsymbol{Z = \frac{\mu_d - \mu_D}{S_d / \sqrt{n}} \sim N(0,1)}$$

ومنه يمكن استنتاج فترة الثقة لمتوسط مجتمع الفروق وهي كما يلي:

$$\boldsymbol{\mu_d - Z_{\frac{\alpha}{2}}\frac{S_d}{\sqrt{n}} < \mu_D < \mu_d + Z_{\frac{\alpha}{2}}\frac{S_d}{\sqrt{n}}}$$

2- في حالة $n < 30$ يكون

$$\boldsymbol{T = \frac{\mu_d - \mu_D}{S_d / \sqrt{n-1}} \sim t(n-1)}$$

و تكون فترة الثقة لمتوسط مجتمع الفروق هي:

$$\boldsymbol{\mu_d - t_{(\frac{\alpha}{2},n-1)}\frac{S_d}{\sqrt{n-1}} < \mu_D < \mu_d + t_{(\frac{\alpha}{2},n-1)}\frac{S_d}{\sqrt{n-1}}}$$

مثال3.13: تم قياس نبض القلب لعينة من 6 رياضيين فكانت كما يلي:

قبل التمرين	72	70	60	65	70	67
بعد التمرين	99	90	80	93	95	90

أوجد فترة ثقة 90% لفرق متوسطي النبض بفرض أن المجتمع طبيعي.

الحل/

قبل التمرين Y_i	72	70	60	65	70	67
بعد التمرين X_i	99	90	80	93	95	90
$D_i = X_i - Y_i$	27	20	20	28	25	23

$$\mu_d = \frac{\sum D_i}{n} = 23.83$$

$$S_d = \sqrt{\frac{\sum(D_i - \mu_d)^2}{n-1}} = 3.43$$

لاحظ أن حجم العينة $n = 6$ وهو أقل من 30 لذلك سنستخدم توزيع t ، وحيث أن $\alpha = 0.1$ نجد من جدول(2) لتوزيع t أن $t_{(0.05,5)} = 2.015$

فتكون فترة الثقة المطلوبة هي:

$$\mu_d - t_{(\frac{\alpha}{2},n-1)}\frac{S_d}{\sqrt{n-1}} < \mu_D < \mu_d + t_{(\frac{\alpha}{2},n-1)}\frac{S_d}{\sqrt{n-1}}$$

$$23.83 - 2.015\frac{3.43}{\sqrt{5}} < \mu_D < 23.83 + 2.015\frac{3.43}{\sqrt{5}}$$

$$20.739 < \mu_D < 26.921$$

رابعاً/ تقدير النسبة بفترة

درسنا في الفصل الثاني توزيع النسبة للعينة المسحوبة من مجتمع ذي الحدين وسنتحدث الآن عن إيجاد فترة الثقة للنسبة عندما يكون حجم العينة $n \geq 30$.

عرَّفنا نسبة عدد النجاحات في العينة بأنها $\hat{p} = \frac{x}{n}$ حيث x عدد النجاحات وقلنا أن هذه النسبة تمثل متغير عشوائي له توزيع احتمالي يسمى توزيع النسبة بوسط حسابي P حيث $\mu_{\hat{p}} = P$ نسبة عدد النجاحات في المجتمع، وبتباين $\sigma^2_{\hat{p}} = \frac{p(1-p)}{n}$.

وتعلمنا أن $\frac{\hat{p}-P}{\sigma_{\hat{p}}} \sim N(0,1)$ ، حيث $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ وقلنا أنه في حالة عدم معرفة P نقدرها بالنسبة $\hat{p}$.

نعرف فترة ثقة $(1-\alpha)\%$ للنسبة بأنها الفترة التي تحقق ما يلي:

$$p_r\left(-Z_{\frac{\alpha}{2}} < \frac{\hat{p}-P}{\sigma_{\hat{p}}} < Z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

وتكون فترة الثقة:

$$-Z_{\frac{\alpha}{2}} < \frac{\hat{p}-P}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} < Z_{\frac{\alpha}{2}}$$

$$\boldsymbol{\hat{p} - Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < P < \hat{p} + Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$$

ويسمى المقدار $\mathbf{d} = \boldsymbol{Z_{\frac{\alpha}{2}}\sqrt{\frac{P(1-P)}{n}}}$ حد $(1-\alpha)\%$ للخطأ في تقدير P و الجدير بالذكر أن P قد تتوفر لدينا من معلومات سابقة ولكنها غالباً ما تكون مجهولة. وفي أسوأ الظروف ممكن اعتبار أن $P = \frac{1}{2}$ حيث تعطي هذه القيمة أكبر خطأ ممكن أن نحصل عليه.

فإذا تم تحديد أكبر خطأ مسموح به في تقدير P وليكن d فإننا نكون قادرين على حساب حجم العينة اللازم لتحقيق ذلك الحد من الخطأ، وبناءً عليه يكون حجم العينة المطلوب هو:

- في حالة معرفة P

$$Z_{\frac{\alpha}{2}}\sqrt{\frac{\mathrm{P}(1-\mathrm{P})}{n}} \le \mathrm{d} \quad \Rightarrow \quad \frac{\mathrm{P}(1-\mathrm{P})}{n} \le \left(\frac{\mathrm{d}}{Z_{\frac{\alpha}{2}}}\right)^2$$

$$n \ge \left(\frac{Z_{\frac{\alpha}{2}}}{\mathrm{d}}\right)^2 \mathrm{P}(1-\mathrm{P})$$

- في حالة عدم معرفة P

$$n \ge \frac{1}{4}\left(\frac{Z_{\frac{\alpha}{2}}}{\mathrm{d}}\right)^2$$

حيث تم التعويض عن $\mathrm{P} = \frac{1}{2}$ وذلك في أسوأ الظروف التي من الممكن أن تقع.

مثال3.14: لإيجاد فترة ثقة 95% لنسبة عدد الطلبة في المدارس الاعدادية الذين يستعملون النظارات الطبية، أخذت عينة عشوائية حجمها 900 طالب فوجد أن عدد مستعملي النظارات الطبية 100 طالب، أوجد فترة الثقة المطلوبة.

الحل/ حيث أن $\alpha = 0.05$ إذن $Z_{\frac{\alpha}{2}} = 1.96$ ولاحظ أن نسبة عدد من يستعمل النظارات في المدارس الاعدادية مجهول لذلك سنقدره بنسبة العينة كما يلي:

$$\hat{\mathrm{p}} = \frac{x}{n} = \frac{100}{900} = \frac{1}{9}$$

فتكون فترة الثقة المطلوبة هي:

$$\hat{\mathrm{p}} - Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{\mathrm{p}}(1-\hat{\mathrm{p}})}{n}} < \mathrm{P} < \hat{\mathrm{p}} + Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{\mathrm{p}}(1-\hat{\mathrm{p}})}{n}}$$

$$\frac{1}{9} - 1.96\sqrt{\frac{\frac{1}{9}\times\frac{8}{9}}{900}} < \mathrm{P} < \frac{1}{9} + 1.96\sqrt{\frac{\frac{1}{9}\times\frac{8}{9}}{900}}$$

$$0.091 < \mathrm{P} < 0.131$$

مثال3.15: في إحدى تجارب علم النفس، يسمح للأشخاص الخاضعين لإحدى التجارب بالاستجابة لاحد مؤشرين B, A . ويريد الباحث أن يُقَدِّر نسبة الأشخاص الذين يختارون المؤشر A ولنرمز لهذه النسبة بالرمز P. كم شخصاً يجب أن يخضع لهذه الدراسة كي نكون واثقين بنسبة 90% أن الخطأ في تقدير P لا يزيد عن 0.04 في الحالتين التاليتين:

1- إذا كنا نعلم أن $P = 0.2$.

2- إذا لم يكن لدينا أية فكرة عن P.

الحل/

1- عند مستوى ثقة 90% تكون $Z_{\frac{\alpha}{2}} = 1.64$ فيكون الحجم المطلوب:

$$n \geq \left(\frac{1.64}{0.04}\right)^2 0.2 \times 0.8 = 268.96$$

فنأخذ حجم العينة $n = 269$

2- بالتعويض المباشر نحصل على

$$n \geq \frac{1}{4}\left(\frac{1.64}{0.04}\right)^2 = 420.25$$

فنأخذ حجم العينة $n = 421$

خامساً/ تقدير الفرق بين نسبتين بفترة

نعلم أنه إذا كان $X_1 \sim b(N_1, P_1)$ وسحبت من مجتمعه عينة حجمها $n_1 \geq 30$ و $X_2 \sim b(N_2, P_2)$ وسحبت من مجتمعه عينة حجمها $n_2 \geq 30$ وكان المجتمعين مستقلين فإن:

$$\mu_{\hat{p}_1 - \hat{p}_2} = P_1 - P_2$$

$$\sigma^2_{\hat{p}_1 - \hat{p}_2} = \frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_2}$$

وأن توزيع فرق النسب المعياري

$$\frac{(\hat{p}_1 - \hat{p}_2) - (P_1 - P_2)}{\sigma_{\hat{p}_1 - \hat{p}_2}} \sim N(0,1)$$

في حال مجهولية P_2, P_1 نأخذ $\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

وبالتالي تكون فترة $(1-\alpha)\%$ ثقة للفرق بين النسب هي:

$$(\hat{p}_1 - \hat{p}_2) - Z_{\frac{\alpha}{2}}\sigma_{\hat{p}_1-\hat{p}_2} < P_1 - P_2 < (\hat{p}_1 - \hat{p}_2) + Z_{\frac{\alpha}{2}}\sigma_{\hat{p}_1-\hat{p}_2}$$

مثال3.16: سجلت 80 حالة نجاح عملية في مشفى A من بين 90 عملية و في المشفى B سجلت 50 عملية نجاح من بين 70 عملية. أوجد فترة ثقة 90% للفرق بين نسبتي النجاح في المشفيين.

الحل/

من المسألة نجد أن:

عند مستوى ثقة 90% تكون $Z_{\frac{\alpha}{2}} = 1.64$

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{80}{90} = \frac{8}{9}$$

$$\hat{p}_2 = \frac{x_2}{n_2} = \frac{50}{70} = \frac{5}{7}$$

$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \sqrt{\frac{\frac{8}{9}\times\frac{1}{9}}{90} + \frac{\frac{5}{7}\times\frac{2}{7}}{70}} = 0.063$$

وتكون فترة الثقة

$$(\hat{p}_1 - \hat{p}_2) - Z_{\frac{\alpha}{2}}\sigma_{\hat{p}_1-\hat{p}_2} < P_1 - P_2 < (\hat{p}_1 - \hat{p}_2) + Z_{\frac{\alpha}{2}}\sigma_{\hat{p}_1-\hat{p}_2}$$

$$\left(\frac{8}{9} - \frac{5}{7}\right) - 1.64 \times 0.063 < P_1 - P_2 < \left(\frac{8}{9} - \frac{5}{7}\right) + 1.64 \times 0.063$$

$$0.071 < P_1 - P_2 < 0.278$$

سادساً/ تقدير تباين مجتمع بفترة

من نظرية 2.4 نعلم أن :

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

ونعرف فترة ثقة $(1-\alpha)$% لتباين مجتمع بأنها الفترة التي تحقق:

$$P\left(\chi^2_{(1-\frac{\alpha}{2},n-1)} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{(\frac{\alpha}{2},n-1)}\right) = 1-\alpha$$

بحل المتباينة بالنسبة للمقدار σ^2 **نجد أن فترة الثقة** للتباين كما يلي:

$$\frac{(n-1)S^2}{\chi^2_{(\frac{\alpha}{2},n-1)}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2},n-1)}}$$

مثال3.17 : أوجد فترة ثقة 95% لتباين مجتمع سحبت منه عينة حجمها 6 وتباينها 11.8

الحل/

لدينا من المسألة n = 6 و $S^2 = 11.8$ وحيث أن $\alpha = 0.05$ ، نجد من جدول(3) لتوزيع كاي تربيع أن:

$$\chi^2_{(0.025,5)} = 12.833 \quad , \quad \chi^2_{(0.975,5)} = 0.831$$

فتكون فترة الثقة كما يلي:

$$\frac{(n-1)S^2}{\chi^2_{(\frac{\alpha}{2},n-1)}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2},n-1)}}$$

$$\frac{5 \times 11.8}{12.833} < \sigma^2 < \frac{5 \times 11.8}{0.831}$$

$$4.6 < \sigma^2 < 71$$

سابعاً/ تقدير النسبة بين تبايني مجتمعين بفترة

نعلم من خلال دراستنا في الفصل الثاني أن:

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1-1, n_2-1)$$

نعرف فتكون فترة ثقة $(1-\alpha)\%$ للنسبة $\frac{\sigma_2^2}{\sigma_1^2}$ بأنها الفترة التي تحقق:

$$\boldsymbol{P\left(\frac{1}{F(\frac{\alpha}{2}, n_2-1, n_1-1)} < \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} < F(\frac{\alpha}{2}, n_1-1, n_2-1)\right) = 1-\alpha}$$

فبحل المتباينة بالنسبة للمقدار $\frac{\sigma_2^2}{\sigma_1^2}$ **نجد أن فترة ثقة** $(1-\alpha)\%$ كما يلي:

$$\boldsymbol{\frac{S_2^2}{S_1^2}\frac{1}{F(\frac{\alpha}{2}, n_2-1, n_1-1)} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{S_2^2}{S_1^2}F(\frac{\alpha}{2}, n_1-1, n_2-1)}$$

وتكمن أهمية ايجاد هذه الفترة في بحث وجود تجانس بين المجتمعين، فكلما كانت النسبة قريبة من 1 كان المجتمعان أكثر تجانساً أي لهما نفس التباين تقريباً.

مثال3.18: سحبت عينة من مجتمع طبيعي حجمها 10 بتباين 9 وسحبت عينة أخرى من مجتمع طبيعي آخر حجمها 15 بتباين 8 أوجد فترة ثقة 95% للنسبة بين تبايني المجتمعين.

الحل/

لدينا من المسألة $n_1 = 10$ و $S_1^2 = 9$ وكذلك $n_2 = 15$ و $S_2^2 = 8$

وحيث أن $\alpha = 0.05$ إذن من جدول توزيع F نجد أن:

$$\mathrm{F}\left(\frac{\alpha}{2}, n_1 - 1, n_2 - 1\right) = \mathrm{F}(0.025,9,14) = 3.2093$$

$$\frac{1}{\mathrm{F}(\frac{\alpha}{2}, n_2 - 1, n_1 - 1)} = \frac{1}{\mathrm{F}(0.025,14,9)} = \frac{1}{3.7694}$$

فتكون فترة الثقة المطلوبة:

$$\frac{8}{9}\frac{1}{3.7694} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{8}{9}3.2093$$

$$0.2358 < \frac{\sigma_2^2}{\sigma_1^2} < 2.8527$$

تمارين:

1- كانت محتويات 9 عبوات أدوية من إنتاج مصنع أدوية كالتالي: 10.1، 10.3، 9.9، 9.8، 10.2، 9.7، 10، 9.7، 10.3 كيلو جرام. أوجد فترة 99% ثقة لوسط إنتاج المصنع من العبوات بافتراض أن الانتاج طبيعي التوزيع.

2- اخذت عينة من 50 خيطاً من مصنع للخيوط فوجد أن وسط قوة هذه الخيوط 80.2 كغم بانحراف معياري 6.5 كغم أوجد فترة ثقة 95% لمعدل قوة جميع الخيوط من إنتاج المصنع.

3- أخذت عينة عشوائية حجمها 400 من معلمي المرحلة الثانوية فوجد أن 80 منهم حاصلون على شهادة الماجستير. قدر نسبة المعلمين في المرحلة الاعدادية الحاصلين على شهادة الماجستير.

4- يراد تصميم دراسة طبية لتقدير نسبة المواطنين الذين يعانون من مشاكل في النظر. كم شخصاً يجب فحصهم كي نكون واثقين بنسبة 98% أن الخطأ في تقدير هذه النسبة لا يزيد عن 0.05 في الحالتين التاليتين:

- إذا لم يكن لدينا أية معلومات سابقة عن هذه النسبة.
- إذا كنا نعلم من دراسات سابقة أن هذه النسبة قد تكون حوالي 0.3

5- أوجد فترة 95% ثقة للفرق بين وسطي المجتمعين الطبيعيين المستقلين في تجربة سحب العينة {9، 11، 10، 13، 12، 9} من المجتمع الاول و العينة {9، 10، 6، 7، 6، 9، 8، 9، 10} من المجتمع الثاني.

6- كون فترة ثقة بنسبة 99% لنسبة من يحملون شهادة عليا في مدينة رفح من خلال العينة العشوائية التي بحجم 150 شخص وتبين أن 10 منهم يحملون شهادة عليا.

7- كون فترة 95% ثقة للفرق بين نسبتي الاناث في البلد A و البلد B ، وذلك من خلال عينة عشوائية بحجم 300 مواطن من البلد A كان عدد الاناث فيها 140 وعينة أخرى من البلد B بحجم 400 مواطن بلغ عدد الاناث فيها 220.

8- ما هو مستوى الثقة المعتمد في تجربة كانت فيها فترة الثقة لوسط المجتمع [20,30] وحجم العينة المستخدم 16 مشاهدة من مجتمع طبيعي التوزيع بتباين 81.

9- ما حجم العينة العشوائية اللازم اختيارها كي لا يزيد طول فترة 99% ثقة لتقدير الوسط الحسابي عن 2.8 سم إذا علمت أن الانحراف المعياري هو 1.3سم.

10-أوجد فترة 90% ثقة للنسبة بين تبايني مجتمعين سحبت من الأول عينة حجمها 10 بتباين 4 ومن الثاني عينة حجمها 7 بتباين 9 .

11-اليك الجدول التالي:

العينة	A	B
حجم العينة	100	49
وسط العينة	60	65
تباين العينة	4	9
نسبة العينة	0.3	0.7

أوجد ما يلي:

- فترة ثقة 90% للفرق بين الوسطين.
- فترة 97% ثقة للفرق بين النسبتين.
- فترة 95% للنسبة بين تبايني المجتمعين.

الفصل الرابع

اختبار الفرضيات

Testing Hypotheses

درسنا في مبادئ الاحصاء أن الاحصاء ينقسم لقسمين، الاحصاء الوصفي Descriptive Statistics والاحصاء الاستنتاجي Inferential Statistics و يتفرع الاخير بدروه لفرعين، الأول يعرف بنظرية التقدير و الثاني باختبارات الفرضيات أو اختبارات المعنوية وهذا ما سنتطرق له في هذا الفصل.

فعند الرغبة في التحقق من ادعاءٍ (فرضية) ما أو التحقق من معلومة معينة عن المجتمع الاحصائي يتم عادة سحب عينة عشوائية من هذا المجتمع و حساب بعض المعالم لها ثم نستخدم طرق احصائية معينة للتحقق من صحة أو عدم صحة هذا الادعاء وقبل الخوض في هذه الاختبارات سنذكر بعض المفاهيم الخاصة التي ستمكننا من فهم الاختبارات وتطبيقها.

الفرضية الاحصائية (Statistical Hypotheses): هي كل عبارة تكون صحتها أو عدم صحتها يحتاج إلى قرار.

و يرمز للفرضية بالرمز H فمثلاً لو كانت الفرضية أن متوسط درجات الطلاب يساوي 75 درجة فيمكن أن نصيغ الفرضية كالتالي:

$$H: \mu = 75$$

الفرضية الصفرية(Null Hypotheses): وتسمى فرضية العدم وتصاغ عادةً بحيث تنفي وجود فروق جوهرية بين معالم العينة و معالم المجتمع ومن هنا سميت بفرضية العدم. وهي الفرضية التي يتم اختبار امكانية رفضها على اعتبار انها صحيحة ويرمز لها بالرمز H_0 .

الفرضية البديلة(Alternative Hypotheses): وهي فرضية مكملة لفرضية العدم حيث يتم قبولها عند رفض فرضية العدم أو رفضها عند قبول فرضية العدم و يرمز لها بالرمز H_1 وتأخذ هذه الفرضية أشكالا مختلفة، وتصاغ الفرضية الصفرية كما يلي:

$$H_0: \mu = \mu_0$$

وتصاغ الفرضية البديلة بأحد الطرق التالية:

$$H_1: \mu \neq \mu_0$$
$$H_1: \mu > \mu_0$$
$$H_1: \mu < \mu_0$$

حيث تشير كل صورة للفرضية الصفرية الى جهة الاختبار ما إذا كان من جهة واحدة أو من جهتين. فمثلا الفرضية $H_1: \mu \neq \mu_0$ تعني أن وسط المجتمع لا يساوي μ_0 وإنما اكبر منها أو أصغر منها وهي تعبر عن أن الاختبار ذو جهتين Two Tails، في حين أن الفرضية $H_1: \mu >$ μ_0 تعني أن وسط المجتمع أكبر من μ_0 أما الفرضية $H_1: \mu < \mu_0$ تعني أن وسط المجتمع أصغر من μ_0 وهما تعنيان أن الاختبار من جهة واحدة One Tail.

قد يحدث أحياناً أن نقبل الفرضية الصفرية وهي خاطئة أو أن نرفضها وهي صحيحة وهذا يقودنا لذكر هذين الخطأين بشيء من التوضيح.

الخطأ من النوع الأول Type I Error : يحدث هذا الخطأ عند رفض H_0 وهي صحيحة ويرمز لاحتمال هذا الخطأ بالرمز **α** وتسمى **α** مستوى المعنوية أو مستوى الدلالة، ويحدد هذا المستوى من قبل الباحث حسب الدراسة وحسب الدقة المطلوبة. وتؤخذ قيمة α في الحسابات كاملة إذا كان الاختبار في اتجاه واحد وتقسم لنصفين إذا كان الاختبار في اتجاهين.

وتحسب **α** من العلاقة:

$$\boldsymbol{\alpha} = \boldsymbol{P}(\textbf{reject } \boldsymbol{H_0} |\ \boldsymbol{H_0} \textbf{is true})$$

الخطأ من النوع الثاني Type II Error : يحدث هذا الخطأ عند قبول H_0 وهي خاطئة ولكن نتائج التجربة أكدت قبولها ويرمز لاحتمال هذا الخطأ بالرمز **β** ، ويحسب من العلاقة:

$$\boldsymbol{\beta} = \boldsymbol{P}(\textbf{accept } \boldsymbol{H_0} |\ \boldsymbol{H_0} \textbf{is false})$$

وبصورة مكافئة

$$\boldsymbol{\beta} = \boldsymbol{P}(\textbf{accept } \boldsymbol{H_0} |\ \boldsymbol{H_1} \textbf{is true})$$

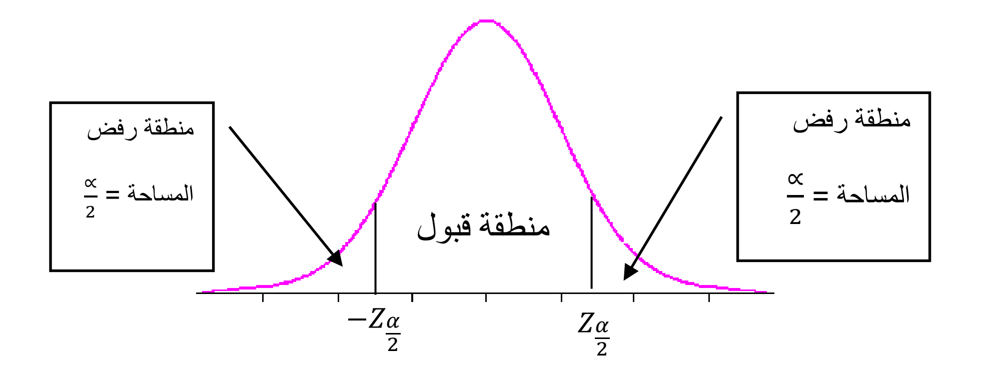
قوة الاختبار(Pot)Power of test: تعرف قوة الاختبار على أنها مقياس لكفاءة الاختبار وبالتالي لدقة الاستنتاج الاحصائي وهي مكملة احتمال الخطأ من النوع الثاني بمعنى:

$$\boldsymbol{Pot = 1 - \beta}$$

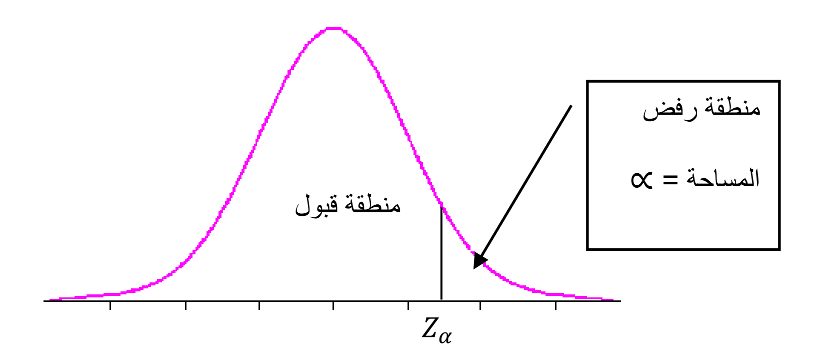
المنطقة الحرجة و القيمة الحرجة المعيارية:

تعرف **المنطقة الحرجة** بأنها المساحة التي تقع أسفل منحنى الاقتران الاحتمالي المستخدم في عملية التحليل الاحصائي وتمثل احتمال رفض H_0 وهي صحيحة، وتسمى هذه المنطقة بمنطقة الرفض وتحدد حسب نوع الفرضية البديلة ويحدد قيمتها مستوى المعنوية α.

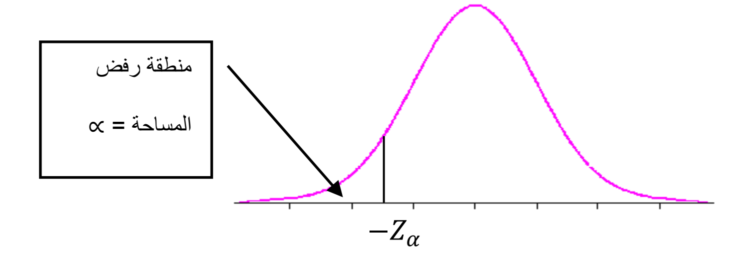
وتسمى باقي المساحة أسفل المنحنى **بمنطقة القبول** وقيم التوزيع الاحتمالي التي تفصل بين المنطقتين **بالقيم الحرجة المعيارية** و الرسم التالي يوضح هذه المفاهيم.



الرسم يوضح مناطق القبول والرفض في حالة $H_1: \mu \neq \mu_0$



الرسم يوضح مناطق القبول والرفض في حالة $H_1: \mu > \mu_0$



الرسم يوضح مناطق القبول والرفض في حالة $H_1: \mu < \mu_0$

بعد أن تعرفنا على بعض المفاهيم الخاصة باختبار الفرضيات نذكر الآن الخطوات المتبعة في عملية اختبار الفرضيات:

1- صياغة الفرضية الصفرية والفرضية البديلة بصورة صحيحة.

2- تحديد مستوى المعنوية المطلوب.

3- تحديد التوزيع المستخدم في الاختبار وحساب القيم الحرجة المعيارية وتسمي هذه القيم أيضاً بالقيم الجدولية.

4- نجد المعيار وهو القيمة التي تحسب من معطيات العينة مستخدمين علاقة خاصة مستنبطة من التوزيع المستخدم.

5- نقارن بين المعيار و القيم الجدولية لتقرير قبول H_0 أم رفضها.

ويتم رفض أو قبول H_0 بطريقتين:

الطريقة الاولى/ مقارنة قيمة المعيار بالقيم الجدولية وذلك كما يلي:

1- إذا كان الاختبار من جهتين أي $H_1: \mu \neq \mu_0$، نرفض H_0 إذا كانت المعيار يقع في منطقة الرفض أي يكون المعيار أكبر من القيمة الجدولية الموجبة أو اصغر من القيمة الجدولية السالبة.

2- إذا كان الاختبار من جهة واحدة نحو اليمين أي $H_1: \mu > \mu_0$ نرفض H_0 إذا كان المعيار أكبر من القيمة الجدولية.

3- إذا كان الاختبار من جهة واحدة نحو اليسار أي $H_1: \mu < \mu_0$ نرفض H_0 إذا كان المعيار اصغر من القيمة الجدولية.

الطريقة الثانية/ المعنوية المحسوبة $\mathbf{p-value}$

تعريف: تعرف المعنوية المحسوبة $\mathbf{p-value}$ بأنها احتمال يحسب لقيمة المعيار، ويؤثر في حسابها التوزيع المستخدم في الاختبار و صيغة الفرضية البديلة وذلك كما يلي:

1- إذا كان الاختبار من جهتين أي $H_1: \mu \neq \mu_0$ فإن:

$$\mathbf{p-value} = \mathbf{2} \times \boldsymbol{P}(\text{المتغير العشوائي} > \text{قيمة المعيار إذا كانت موجبة})$$

أو

$$\mathbf{p-value} = \mathbf{2} \times \boldsymbol{P}(\text{المتغير العشوائي} < \text{قيمة المعيار إذا كانت سالبة})$$

فنرفض H_0 إذا كانت $\mathbf{p-value}$ أقل من مستوى المعنوية ونقبلها إذا كانت $\mathbf{p-value}$ أكبر من أو تساوي مستوى المعنوية.

2- إذا كان الاختبار من جهة واحدة نحو اليمين فإن:

$$\mathbf{p-value} = \boldsymbol{P}(\text{المتغير العشوائي} > \text{قيمة المعيار})$$

فنرفض H_0 إذا كانت $\mathbf{p-value}$ أقل من مستوى المعنوية ونقبلها إذا كانت $\mathbf{p-value}$ أكبر من أو تساوي مستوى المعنوية.

3- إذا كان الاختبار من جهة واحدة نحو اليسار فإن:

$$\mathbf{p-value} = \boldsymbol{P}(\text{المتغير العشوائي} < \text{قيمة المعيار})$$

فنرفض H_0 إذا كانت $\mathbf{p-value}$ أقل من مستوى المعنوية ونقبلها إذا كانت $\mathbf{p-value}$ أكبر من أو تساوي مستوى المعنوية.

نلاحظ مما سبق أن الطريقة الثانية عملية أكثر من الطريقة الاولى حيث أنه في كل حالات الفرضية البديلة يتم قبول الفرضية الصفرية إذا كانت $\mathbf{p-value} \geq \boldsymbol{\alpha}$ وترفض الفرضية الصفرية إذا كانت $\mathbf{p-value} < \boldsymbol{\alpha}$.

والآن سنتطرق لاختبارات مهمة في الاحصاء الاستدلالي يشيع استخدامها في التطبيقات العملية.

أولاً/ اختبارات الوسط الحسابي لمجتمع طبيعي التوزيع.

سنميز هنا حالتين:

- **إذا كان تباين المجتمع معلوم.**

في هذه الحالة يستخدم التوزيع الطبيعي في الاختبار ويكون المعيار في هذه الحالة:

$$\boldsymbol{Z} = \frac{\overline{\boldsymbol{X}} - \boldsymbol{\mu_0}}{\boldsymbol{\sigma}/\sqrt{\boldsymbol{n}}} \sim \boldsymbol{N}(\mathbf{0}, \mathbf{1})$$

والامثلة التالية ستوضح كيفية تطبيق الاختبار بخطوات متتابعة ثابتة.

مثال4.1: سحبت عينة عشوائية من مجتمع طبيعي تباينه 25 وكان حجم العينة 64 ووسطها الحسابي 20، اختبر الفرضية القائلة بأن الوسط الحسابي للمجتمع هو 18 عند مستوى دلالة 0.05

الحل/

<u>الفرضيات:</u>

$$H_0: \mu = 18$$
$$H_1: \mu \neq 18$$

<u>التوزيع والقيم الحرجة المعيارية:</u>

حيث أن مستوى الدلالة $\alpha = 0.05$ إذن $\frac{\alpha}{2} = 0.025$ ، وواضح أن الاختبار من اتجاهين لذلك تكون القيم الحرجة هي:

$$-Z_{\frac{\alpha}{2}} = -1.96 \qquad , \qquad Z_{\frac{\alpha}{2}} = 1.96$$

المعيار:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{20 - 18}{5/8} = 3.2$$

المقارنة و القرار:

الطريقة الاولى:

حيث أن $(Z = 3.2) > (Z_{\frac{\alpha}{2}} = 1.96)$، فإننا نرفض الفرضية الصفرية وبالتالي نقبل الفرضية البديلة ونستنتج أن وسط المجتمع لا يساوي 18

الطريقة الثانية:

بحساب المعنوية المحسوبة ومقارنتها مع مستوى المعنوية المعطى نجد أن:

$$\text{p} - \text{value} = 2 \times P(Z > 3.2) = 0.0014 < \alpha = 0.05$$

وبالتالي فإننا نرفض الفرضية الصفرية.

ملاحظات هامة:

1- بحساب فترة الثقة لوسط المجتمع في المثال السابق نجد أنها [18.775,21.225] ولاحظ أن هذه الفترة لا تحتوي على القيمة التي نختبر صحتها كوسط للمجتمع وهي $\mu_0 = 18$ وبالتالي ففترة الثقة تخبرنا بأنه باحتمال 95% لن تكون هذه القيمة من القيم التي يمكن أن يأخذها وسط المجتمع لذلك نرفض الفرضية الصفرية، وبناءً عليه يمكن حساب فترات الثقة واستخدامها لاتخاذ القرار في مسائل اختبار الفرضيات.

2- لإيجاد القيم الحرجة للمتغير العشوائي $\bar{X}$ نحل المتباينة $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} > Z_{\frac{\alpha}{2}}$ بالنسبة للمتغير $\bar{X}$ فنجد أن الحل هو $\bar{X} > Z_{\frac{\alpha}{2}} \times \sigma/\sqrt{n} + \mu_0$ ، ومن معطيات السؤال نجد أن $\bar{X} > 19.225$ وبالتالي تكون 19.225 هي القيمة الحرجة الاولي ومن ناحية اليمين وبحل المتباينة $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < -Z_{\frac{\alpha}{2}}$ بالنسبة لنفس المتغير نجد أن $\bar{X} < -Z_{\frac{\alpha}{2}} \times \sigma/\sqrt{n} + \mu_0$ ومن معطيات السؤال نجد أن $\bar{X} < 16.775$ وبالتالي تكون 16.775 هي القيمة الحرجة الثانية ومن ناحية اليسار.

3- تستخدم القيم الحرجة السابقة في ايجاد احتمال الوقوع في الخطأ من النوع الاول واحتمال الوقوع في الخطأ من النوع الثاني وبالتالي في حساب قوة الاختبار وذلك كما يلي:

- بمعلومية القيم الحرجة التي حسبناها في الملاحظة الثانية نجد أن:

$$\alpha = P(\text{reject } H_0 |\ H_0 \text{is true})$$
$$\alpha = P(\bar{X} > 19.225 \text{ or } \bar{X} < 16.775 | \mu = 18)$$
$$\alpha = P(\bar{X} > 19.225 | \mu = 18) + P(\bar{X} < 16.775 | \mu = 18)$$
$$\alpha = P\left(Z > \frac{19.225 - 18}{5/8}\right) + P\left(Z < \frac{16.775 - 18}{5/8}\right)$$
$$\alpha = P(Z > 1.96) + P(Z < -1.96) = 0.05$$

ولاحظ أنه نفس مستوى المعنوية الذي استخدمناه في حل المثال السابق.

- وايضاً بمعلومية القيم الحرجة التي حسبناها في الملاحظة الثانية نجد أن احتمال الوقوع في خطأ من النوع الثاني عندما تكون $\mu = 19$ على سبيل المثال حيث تكون $H_1 : \mu = 19$ كما يلي:

$$\beta = P(\text{accept } H_0 |\ H_1 \text{is true})$$
$$\beta = P(16.775 < \bar{X} < 19.225\ | \mu = 19)$$
$$\beta = P\left(\frac{16.775 - 19}{5/8} < Z < \frac{19.225 - 19}{5/8}\right)$$
$$= P(-3.56 < Z < 0.36) = 0.6404$$

وبالتالي تكون قوة الاختبار في هذه الحالة :

$$Pot = 1 - \beta = 1 - 0.6404 = 0.3596$$

4- لاحظ كبر الخطأ من النوع الثاني في الحالة المذكورة وبالتالي ضعف قوة الاختبار والسبب في ذلك قرب قيمة $\mu = 19$ منها في الفرضة الصفرية $H_0 : \mu = 18$ وبالتالي يصعب على الاختبار أن يميز بينهما.

فلو أردنا حساب قوة الاختبار في حالة $H_1 : \mu = 21$ نجد أن:

$$\beta = P(16.775 < \bar{X} < 19.225\ | \mu = 25)$$
$$\beta = P\left(\frac{16.775 - 21}{5/8} < Z < \frac{19.225 - 21}{5/8}\right)$$
$$= P(-6.76 < Z < -2.84) = 0.0021$$

وبالتالي تكون قوة الاختبار في هذه الحالة :

$$Pot = 1 - \beta = 1 - 0.0021 = 0.9979$$

لاحظ في هذه الحالة صغر الخطأ من النوع الثاني وقوة الاختبار في اتخاذ القرار السليم.

مثال4.2: سحبت عينة عشوائية من مجتمع طبيعي التوزيع بانحراف معياري 3 يمثل أوزان تلاميذ في الصف الاول الاساسي وكانت أوزان التلاميذ في العينة هي:

$$\{28,22,23,20,30,29,26,29,26,27\}$$

اختبر عند مستوى معنوية $\alpha = 0.05$ الفرضية $H_0: \mu = 24$ مقابل $H_1: \mu > 24$.

الحل/

من معطيات السؤال نجد أن الوسط الحسابي للعينة هو:

$$\bar{x} = \frac{\sum x}{n} = \frac{260}{10} = 26$$

<u>الفرضيات:</u>

$$H_0: \mu = 24$$
$$H_1: \mu > 24$$

<u>التوزيع والقيم الحرجة المعيارية:</u>

حيث أن مستوى الدلالة $\alpha = 0.05$ وحيث أن الاختبار من اتجاه واحد نحو اليمين إذن تكون القيمة الحرجة هي:

$$Z_\alpha = 1.65$$

<u>المعيار:</u>

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{26 - 24}{3/\sqrt{10}} = 2.11$$

<u>المقارنة و القرار:</u>

الطريقة الاولى:

حيث أن $(Z_{\frac{\alpha}{2}} = 1.65) < (Z = 2.11)$ فإننا نرفض الفرضية الصفرية وبالتالي نقبل الفرضية البديلة وهذا يعني أن معدل أوزان تلاميذ الصف الاول الاساسي أقل من 24 .

الطريقة الثانية:

$$\text{p} - \text{value} = P(Z > 2.11) = 0.0495 < 0.05$$

لذلك فإننا نرفض الفرضية الصفرية.

- **إذا كان تباين المجتمع مجهول.**

1− إذا كان حجم العينة $n \geq 30$

في هذه الحالة يستخدم التوزيع الطبيعي في الاختبار ويكون المعيار في هذه الحالة:

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim N(0,1)$$

مثال4.3: عينة عشوائية بحجم 40 مشاهدة بلغ وسطها الحسابي 12 وتباينها 2.56 سحبت من مجتمع بتوزيع طبيعي فهل يمكننا القول أن متوسط المجتمع يساوي13. أجب عند مستوى معنوية 0.01

الحل/

الفرضيات:

$$H_0: \mu = 13$$
$$H_1: \mu \neq 13$$

التوزيع والقيم الحرجة المعيارية:

حيث أن مستوى الدلالة $\alpha = 0.01$ إذن $\frac{\alpha}{2} = 0.005$ ، وواضح أن الاختبار من اتجاهين لذلك تكون القيم الحرجة هي:

$$-Z_{\frac{\alpha}{2}} = -2.56 \qquad , \qquad Z_{\frac{\alpha}{2}} = 2.56$$

المعيار:

$$Z = \frac{12 - 13}{1.6/\sqrt{40}} = -3.95$$

المقارنة و القرار:

الطريقة الاولى:

حيث أن $(-Z_{\frac{\alpha}{2}} = -2.56) > (Z = -3.95)$ فإننا نرفض الفرضية الصفرية وبالتالي نقبل الفرضية البديلة .

الطريقة الثانية:

$$\text{p} - \text{value} = 2 \times P(Z < -3.95) = 0.0004 < 0.01$$

لذلك فإننا نرفض الفرضية الصفرية.

2- إذا كان حجم العينة $n < 30$

في هذه الحالة يستخدم توزيع t في الاختبار ويكون المعيار في هذه الحالة:

$$\boldsymbol{T = \frac{\bar{X} - \mu_0}{\mathbf{S}/\sqrt{\mathbf{n-1}}} \sim \mathbf{t}(n-1)}$$

مثال4.4: أعد حل مثال4.2 مفترضاً مجهولية تباين المجتمع.

الحل/ حيث أن حجم العينة صغير فالتوزيع المستخدم هو توزيع t بدرجة حرية 9 وبحساب تباين العينة نجد أنه:

$$S = \sqrt{\frac{\sum(x-26)^2}{9}} = 3.33$$

الفرضيات:

$$H_0: \mu = 24$$
$$H_1: \mu > 24$$

التوزيع والقيم الحرجة المعيارية:

وحيث أن مستوى الدلالة $\alpha = 0.05$ وحيث أن الاختبار من اتجاه واحد نحو اليمين إذن تكون القيمة الحرجة هي:

$$t_{(0.05,9)} = 1.833$$

المعيار:

$$T = \frac{26 - 24}{3.33/\sqrt{9}} = 1.802$$

المقارنة و القرار:

الطريقة الاولى:

حيث أن $(\mathbf{T = 1.802}) < (\mathbf{t_{(0.05,9)} = 1.833})$ فإننا نقبل الفرضية الصفرية وهذا يعني أن معدل أوزان تلاميذ الصف الأول الأساسي يساوي 24 .

الطريقة الثانية:

$$\mathrm{p-value} = P(t > 1.802) = 0.05 = \alpha$$

لذلك فإننا نقبل الفرضية الصفرية.

ملاحظة/ في حالة عدم معرفة توزيع المجتمع نلجأ لسحب عينة كبيرة بقدرٍ كافٍ حتى نستطيع تطبيق نظرية النهاية المركزية ونستخدم التوزيع الطبيعي.

ثانياً/ اختبارات الفروق بين متوسطي مجتمعين مستقلين طبيعيي التوزيع.

إذا كان $X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$ وسحبت من مجتمعه عينة حجمها n_1 و $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$ وسحبت من مجتمعه عينة حجمها n_2 وكان المجتمعين مستقلين ونريد اختبار تساوي متوسطي المجتمعين فتكون الفرضية الصفرية كالتالي:

$$H_0 : \mu_1 = \mu_2$$

ويمكن لهذه الفرضية أن تأخذ شكلاً آخر أكثر مرونة وشيوعاً وهو:

$$H_0 : \mu_1 - \mu_2 = 0$$

وتأخذ الفرضية البديلة إحدى الصور:

$$H_1 : \mu_1 - \mu_2 \neq 0$$
$$H_1 : \mu_1 - \mu_2 < 0$$
$$H_1 : \mu_1 - \mu_2 > 0$$

وسنذكر حالات مختلفة لاختبارات الفروق بين متوسطي مجتمعين وهي:

- **عند معلومية σ_1^2 و σ_2^2**

في هذه ا لحالة يستخدم التوزيع الطبيعي في الاختبار ويكون المعيار في هذه الحالة:

$$\mathbf{Z} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim \mathbf{N}(\mathbf{0}, \mathbf{1})$$

مثال4.5: إذا علم أن $X_1 \sim N(\mu_1, 36)$ وسحبت منه عينة حجمها 16 مشاهدة وسطها الحسابي 30 و $X_2 \sim N(\mu_2, 25)$ وسحبت منه عينة حجمها 12 مشاهدة وسطها الحسابي 33 ، اختبر وجود فرق معنوي بين متوسطي المجتمعين عند دلالة 10%

الحل/

من المسألة نجد أن: $n_1 = 16 \ , \ \bar{x}_1 = 30 \ , \ n_1 = 12 \ , \ \bar{x}_2 = 33$

<u>الفرضيات:</u>

$$H_0: \mu_1 - \mu_2 = 0$$
$$H_1: \mu_1 - \mu_2 \neq 0$$

<u>التوزيع و القيم الحرجة المعيارية:</u>

التوزيع المستخدم هو التوزيع الطبيعي وحيث أن $\alpha = 0.1$ والاختبار ذو اتجاهين إذن نحسب $\frac{\alpha}{2} = 0.05$ وتكون القيم الحرجة:

$$-Z_{\frac{\alpha}{2}} = -1.64 \quad ، Z_{\frac{\alpha}{2}} = 1.64$$

<u>المعيار:</u>

$$Z = \frac{30 - 33}{\sqrt{\frac{36}{16} + \frac{25}{12}}} = -1.44$$

المقارنة والقرار:

الطريقة الاولى:

وحيث أن $(Z = -1.44) > (-Z_{\frac{\alpha}{2}} = -1.64)$ فإننا نقبل الفرضية الصفرية وبالتالي المجتمعان لهما تقريباً نفس الوسط الحسابي .

الطريقة الثانية:

$$\mathrm{p-value} = 2 \times P(Z < -1.44) = 0.1498 > 0.1$$

و بالتالي فإننا نقبل الفرضية الصفرية.

ملاحظة هامة:

بحساب فترة الثقة لوسط المجتمع في المثال السابق نجد أنها $[-6.4139, 0.4139]$ ولاحظ أن هذه الفترة تحتوي على القيمة صفر وهي التي نختبر صحتها كفرق بين وسطي المجتمعين وبالتالي ففترة الثقة تخبرنا بأنه باحتمال 99% الفرق بين وسطي المجتمعين يساوي صفر أي يكون وسطي المجتمعين متساوٍ.

- **عند مجهولية σ_1^2 و σ_2^2 وحجم العينتين كبير**

مثال4.6: أخذت عينتان من مجتمعين طبيعيين مستقلين والجدول التالي يبين بعض احصاءاتها.

	المجتمع الأول	المجتمع الثاني
حجم العينة	50	60
متوسط العينة	57.5	54.4
الانحراف المعياري للعينة	6.2	10.6

يدعي أحد الباحثين أن متوسط المجتمع الاول أكبر من متوسط المجتمع الثاني. اختبر صحة هذا الادعاء عند مستوى معنوية 0.05

الحل/

الفرضيات:

$$H_0: \mu_1 - \mu_2 = 0$$
$$H_1: \mu_1 - \mu_2 > 0$$

<u>التوزيع و القيم الحرجة المعيارية:</u>

التوزيع المستخدم هو التوزيع الطبيعي وذلك لكبر حجوم العينتين وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاه واحد نحو اليمين إذن تكون القيمة الحرجة:

$$Z_{\alpha} = 1.64$$

<u>المعيار:</u>

$$\text{Z} = \frac{57.5 - 54.4}{\sqrt{\frac{(6.2)^2}{50} + \frac{(10.6)^2}{60}}} = 1.91$$

<u>المقارنة والقرار:</u>

الطريقة الاولى:

وحيث أن $(Z = 1.91) > (Z_{\alpha} = 1.64)$ فإننا نرفض الفرضية الصفرية وعليه نعتبر أن هذا دليلاً كافياً على أن متوسط المجتمع الاول أكبر من متوسط المجتمع الثاني.

الطريقة الثانية:

$$\text{p} - \text{value} = P(Z > 1.91) = 0.0281 < 0.05$$

و بالتالي فإننا نرفض الفرضية الصفرية.

ملاحظة هامة:

بحساب فترة الثقة لوسط المجتمع في مثال 4.6 نجد أنها [0.435,5.765] ولاحظ أن هذه الفترة لا تحتوي على القيمة صفر وهي التي نختبر صحتها كفرق بين وسطي المجتمعين وبالتالي ففترة الثقة تخبرنا بأنه باحتمال 95% لن يكون الفرق بين وسطي المجتمعين يساوي صفر أي يكون وسطي المجتمعين غير متساوٍ.

وهنا نتساءل، أي المجتمعين له وسط أكبر من الثاني؟

للإجابة على هذا التساؤل لاحظ أننا قبلنا الفرضية القائلة أن $\mu_1 - \mu_2 > 0$ وبالتالي $\mu_1 > \mu_2$ مما يدل على أن المجتمع الاول له وسط حسابي اكبر من المجتمع الثاني.

- **عند مجهولية تباين المجتمعين الطبيعيين وكان على الأقل حجم أحد العينتين صغير**

في هذه الحالة يستخدم توزيع t في الاختبار وسنميز هنا حالتين:

3- **إذا كان** $\boldsymbol{\sigma_1^2 = \sigma_2^2 = \sigma^2}$

يكون المعيار في هذه الحالة:

$$\boldsymbol{T} = \frac{\boldsymbol{\bar{X}_1 - \bar{X}_2}}{\boldsymbol{S_{\bar{X}_1-\bar{X}_2}}} \sim \boldsymbol{t(n_1 + n_2 - 2)}$$

حيث

$$S^2_{\bar{X}_1-\bar{X}_2} = S^2_p\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

$$S^2_p = \frac{(n_1 - 1)S^2_1 + (n_2 - 1)S^2_2}{n_1 + n_2 - 2}$$

مثال4.7: إذا علم أن $X_1 \sim \mathrm{N}(\mu_1, \sigma^2)$ وسحبت منه العينة {5،6،3،7،8} و $X_2 \sim \mathrm{N}(\mu_2, \sigma^2)$ وسحبت منه العينة {12،10،8،12،9،5} ، اختبر وجود فرق معنوي بين متوسطي المجتمعين عند مستوى دلالة 5%.

الحل/

من معطيات السؤال نحسب ما يلي:

$$\bar{x}_1 = 5.8 \quad , \quad \bar{x}_2 = 9.33 \quad , \quad S^2_1 = 3.7 \quad , \quad S^2_2 = 7.076$$

$$S^2_p = \frac{(5-1) \times 3.7 + (6-1) \times 7.076}{5 + 6 - 2} = 5.58$$

$$S^2_{\bar{X}_1-\bar{X}_2} = 5.58 \times \left(\frac{1}{5} + \frac{1}{6}\right) = 2.046$$

<u>الفرضيات:</u>

$$H_0: \mu_1 - \mu_2 = 0$$
$$H_1: \mu_1 - \mu_2 \neq 0$$

<u>التوزيع و القيم الحرجة المعيارية:</u>

التوزيع المستخدم هو توزيع t بدرجات حرية 9 وذلك لصغر حجوم العينتين وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاهين إذن تكون القيم الحرجة:

$$-t_{(0.025,9)} = -2.262 \quad ، \quad t_{(0.025,9)} = 2.262$$

المعيار:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}} = \frac{5.8 - 9.33}{1.43} = -2.47$$

المقارنة والقرار:

الطريقة الاولى:

وحيث أن $(T = -2.47) < (-t_{(0.025,9)} = -2.262)$ فإننا نرفض الفرضية الصفرية وعليه نعتبر انه يوجد فرق معنوي بين متوسطي المجتمعين.

الطريقة الثانية:

$$\text{p} - \text{value} = 2 \times P(t < -2.47) = 0.02 < 0.1$$

و بالتالي فإننا نرفض الفرضية الصفرية.

4- عند عدم تساوي تبايني المجتمعين

يكون المعيار في هذه الحالة:

$$\mathbf{T} = \frac{\bar{\boldsymbol{X}}_1 - \bar{\boldsymbol{X}}_2}{\boldsymbol{S}_{\bar{X}_1 - \bar{X}_2}} \sim \mathbf{t(f)}$$

حيث

$$\boldsymbol{S}^2_{\bar{X}_1 - \bar{X}_2} = \frac{\boldsymbol{S}_1^2}{\boldsymbol{n}_1} + \frac{\boldsymbol{S}_2^2}{\boldsymbol{n}_2}$$

وتحسب درجة الحرية من العلاقة:

$$\mathbf{f} = \frac{\left[\frac{\boldsymbol{S}_1^2}{\boldsymbol{n}_1} + \frac{\boldsymbol{S}_2^2}{\boldsymbol{n}_2}\right]^2}{\frac{\left({\boldsymbol{S}_1^2}/{\boldsymbol{n}_1}\right)^2}{\boldsymbol{n}_1 - \mathbf{1}} + \frac{\left({\boldsymbol{S}_2^2}/{\boldsymbol{n}_2}\right)^2}{\boldsymbol{n}_2 - \mathbf{1}}}$$

مثال4.8: أعد حل مثال4.7 بفرض عدم تساوي تبايني المجتمعين.

الحل/

الفرضيات:

$$H_0: \mu_1 - \mu_2 = 0$$
$$H_1: \mu_1 - \mu_2 \neq 0$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع t وذلك لصغر حجوم العينتين وتكون درجة الحرية :

$$\bar{x}_1 = 5.8 \quad , \quad \bar{x}_2 = 9.33 \quad , \quad S_1^2 = 3.7 \quad , \quad S_2^2 = 7.076$$

$$\mathrm{f} = \frac{\left[\frac{3.7}{5} + \frac{7.076}{6}\right]^2}{\frac{\left({}^{3.7}/_{5}\right)^2}{4} + \frac{\left({}^{7.076}/_{6}\right)^2}{5}} = \frac{1.92}{0.137 + 0.278} = 8.88$$

فنأخذ درجة الحرية تساوي 9

وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاهين إذن تكون القيم الحرجة:

$$-t_{(0.025,9)} = -2.262 \quad ، \quad t_{(0.025,9)} = 2.262$$

المعيار:

$$S_{\bar{X}_1-\bar{X}_2}^2 = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} = \frac{3.702}{5} + \frac{7.076}{6} = 1.92$$

$$\mathrm{T} = \frac{5.8 - 9.33}{1.39} = -2.54$$

المقارنة والقرار:

الطريقة الاولى:

حيث أن $(-t_{(0.025,9)} = -2.262) < (\mathrm{T} = -2.54)$ فإننا نرفض الفرضية الصفرية وعليه نعتبر انه يوجد فرق معنوي بين متوسطي المجتمعين.

الطريقة الثانية:

$$\mathrm{p-value} = 2 \times P(t < -2.54) = 0.02 < 0.05$$

و بالتالي فإننا نرفض الفرضية الصفرية.

ثالثاً/ اختبارات الفروق بين متوسطي مجتمعين مرتبطين طبيعيي التوزيع

بفرض أن $X\sim N(\mu_1,\sigma_1^2)$ و $Y\sim N(\mu_2,\sigma_2^2)$ وكان المجتمعان مرتبطين وسحبت من كل مجتمع عينة حجمها n $\{X_1,X_2,\ldots\ldots,X_n\}$ و $\{Y_1,Y_2,\ldots\ldots,Y_n\}$ على الترتيب، بحيث تمثل X_i و Y_i القيمتين المتناظرتين في العينتين، إذا كانت الفروق بين قيم العينتين المتناظرة هي $\{D_1,D_2,\ldots\ldots,D_n\}$ حيث $D_i = X_i - Y_i$ لجميع قيم $i = 1,2,\ldots\ldots,n$

إذا رغبنا في اختبار تساوي وسطي المجتمعين فمن المنطق أن نصيغ الفرضية الصفرية كالتالي:

$$H_0: \mu_D = 0$$

وتأخذ الفرضية الصفرية احدى الصور التالية:

$$H_1: \mu_D \neq 0$$
$$H_1: \mu_D < 0$$
$$H_1: \mu_D > 0$$

وسنميز هنا حالتين:

- الحالة الأولى عندما تكون $\boldsymbol{n \geq 30}$

يكون المعيار في هذه الحالة:

$$\boldsymbol{Z = \frac{\mu_d}{S_d / \sqrt{n}} \sim N(0,1)}$$

مثال4.9: إذا كان $X\sim N(\mu_1,\sigma_1^2)$ و $Y\sim N(\mu_2,\sigma_2^2)$ وكان المجتمعان مرتبطين وسحبت من كل مجتمع عينة حجمها 30 وكان مجموع الفروق بين القيم المتناظرة للعينتين هو 180 و الانحراف المعياري للفروق 9 فهل يمكن القول أن المجتمعين لهما نفس المتوسط عند مستوى معنوية 0.05

الحل/

الفرضيات:

$$H_0: \mu_D = 0$$
$$H_1: \mu_D \neq 0$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع Z وذلك لكبر حجم العينة وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاهين إذن تكون القيم الحرجة:

$$-Z_{\frac{\alpha}{2}} = -1.96 \quad ، \quad Z_{\frac{\alpha}{2}} = 1.96$$

المعيار:

$$\mu_d = \frac{\sum D_i}{n} = \frac{180}{30} = 6$$
$$Z = \frac{\mu_d}{S_d / \sqrt{n}} = \frac{6}{9/\sqrt{30}} = 3.65$$

المقارنة والقرار:

الطريقة الاولى:

وحيث أن $(Z = 3.65) > (Z_{\frac{\alpha}{2}} = 1.96)$ فإننا نرفض الفرضية الصفرية وعليه نعتبر انه يوجد فرق معنوي بين متوسطي المجتمعين.

الطريقة الثانية:

$$\text{p} - \text{value} = 2 \times P(Z > 3.65) = 0.0004 < 0.05$$

و بالتالي فإننا نرفض الفرضية الصفرية.

- الحالة الثانية عندما تكون $\boldsymbol{n < 30}$

يكون المعيار في هذه الحالة:

$$\mathbf{T} = \frac{\boldsymbol{\mu_d}}{\boldsymbol{S_d} / \sqrt{\boldsymbol{n-1}}} \sim \boldsymbol{t(n-1)}$$

مثال4.10: لديك بيانات تمثل إنجاز سباق 100م لطلبة الكلية قبل وبعد أخذ دورة تدريبية في السباق.

بعد	15	16	16	17	14	18
قبل	14	14	13	14	12	14

فهل هناك فرق جوهري بين أداء الطلبة قبل وبعد الدورة، اختبر ذلك عند مستوى معنوية 0.05

الحل/

بعد X_i	15	16	16	17	14	18
قبل Y_i	14	14	13	14	12	14
$D_i = X_i - Y_i$	1	2	3	3	2	4

$$\mu_d = \frac{\sum D_i}{n} = \frac{15}{6} = 2.5$$

$$S_d = \sqrt{\frac{\sum(D_i - \mu_d)^2}{n-1}} = 1.05$$

الفرضيات:

$$H_0: \mu_D = 0$$
$$H_1: \mu_D \neq 0$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع t بدرجة حرية 5 وذلك لصغر حجم العينة وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاهين إذن تكون القيم الحرجة:

$$-t_{(0.025,5)} = -2.571 \quad ، \quad t_{(0.025,5)} = 2.571$$

المعيار:

$$\mathrm{T} = \frac{\mu_d}{S_d/\sqrt{n-1}} = \frac{2.5}{1.05/\sqrt{5}} = 5.32$$

المقارنة والقرار:

الطريقة الاولى:

حيث أن $(T = 5.32) > (t_{(0.025,5)} = 2.571)$ فإننا نرفض الفرضية الصفرية وعليه نعتبر أنه يوجد فرق معنوي بين الاداء قبل وبعد الدورة.

الطريقة الثانية:

$$\mathrm{p-value} = 2 \times P(t > 5.32) = 0.002 < 0.05$$

و بالتالي فإننا نرفض الفرضية الصفرية.

ملاحظة/ إن صياغة الفرضية البديلة يحددها الهدف من التجربة فمثلاً في تجربة لمعرفة تأثير دواء جديد لمعالجة مرض السكر إذا كان تساؤلنا:

هل هناك تحسناً بعد تناول الدواء الجديد؟ تكون الفرضية البديلة $H_1: \mu_D > 0$

هل هناك تردياً بعد تناول الدواء الجديد؟ تكون الفرضية البديلة $H_1: \mu_D < 0$

هل هناك فرقاً في القياسات عند تناول الدواء الجديد؟ تكون الفرضية البديلة $H_1: \mu_D \neq 0$

ثالثاً/ اختبار الفرضيات المتعلقة بالنسبة.

تأخذ الفرضية الصفرية الشكل:

$$H_0: \mathrm{P} = \mathrm{P}_0$$

والفرضية البديلة أحد الصور:

$$H_1: \mathrm{P} \neq \mathrm{P}_0$$
$$H_1: \mathrm{P} > \mathrm{P}_0$$
$$H_1: \mathrm{P} < \mathrm{P}_0$$

ويكون المعيار في هذه الحالة:

$$\boldsymbol{Z} = \frac{\widehat{\mathbf{P}} - \mathbf{P_0}}{\sqrt{\dfrac{\mathbf{P_0}(\mathbf{1} - \mathbf{P_0})}{\boldsymbol{n}}}} \sim \mathbf{N}(\mathbf{0}, \mathbf{1})$$

مثال4.11: إذا كانت نسبة مستعملي حزام الأمان في السيارات (قبل تشريع الزام الاستعمال) هي 0.8 ، درست عينة عشوائية حجمها 200 سائق بعد صدور تشريع الالزام فوجد أن 170 منهم يستعملون الحزام ، اختبر على مستوى دلالة 0.05 ما إذا كان التشريع قد زاد نسبة المستعملين له.

الحل/ واضح من المسألة أن $\mathrm{P} = 0.8$ وأن $\widehat{\mathrm{P}} = \frac{170}{200} = 0.85$.

الفرضيات:

$$H_0: \mathrm{P} = 0.8$$
$$H_1: \mathrm{P} > 0.8$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو التوزيع الطبيعي وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاه واحد نحو اليمين إذن تكون القيم الحرجة:

$$Z_\alpha = 1.64$$

المعيار:

$$Z = \frac{0.85 - 0.8}{\sqrt{\frac{0.8 \times 0.2}{200}}} = 1.8$$

المقارنة والقرار:

الطريقة الاولى:

وحيث أن $(Z = 1.8) > (Z_\alpha = 1.64)$ فإننا نرفض الفرضية الصفرية وبالتالي فإن صدور التشريع قد زاد من نسبة مستعملي حزام الأمان .

الطريقة الثانية:

$$\mathrm{p-value} = P(Z > 1.8) = 0.0359 < 0.05$$

و بالتالي فإننا نرفض الفرضية الصفرية.

رابعاً/ اختبار الفرق بين نسبتين.

بفرض أن $X_1 \sim \mathrm{b}(\mathrm{N}_1, \mathrm{P}_1)$ وسحبت من مجتمعه عينة حجمها $n_1 \geq 30$ و $X_2 \sim \mathrm{b}(\mathrm{N}_2, \mathrm{P}_2)$ وسحبت من مجتمعه عينة حجمها $n_2 \geq 30$ وكان المجتمعان مستقلين ونريد اختبار تساوي نسب المجتمعين فتكون الفرضية الصفرية كالتالي:

$$H_0: \mathrm{P}_1 - \mathrm{P}_2 = 0$$

وتأخذ الفرضية البديلة إحدى الصور:

$$H_1: \mathrm{P}_1 - \mathrm{P}_2 \neq 0$$
$$H_1: \mathrm{P}_1 - \mathrm{P}_2 < 0$$
$$H_1: \mathrm{P}_1 - \mathrm{P}_2 > 0$$

ويكون المعيار المستخدم هو:

$$\mathbf{Z} = \frac{\hat{\mathbf{p}}_1 - \hat{\mathbf{p}}_2}{\sqrt{\frac{\mathbf{p}_1(1-\mathbf{p}_1)}{n_1} + \frac{\mathbf{p}_2(1-\mathbf{p}_2)}{n_2}}} \sim \mathbf{N}(0,1)$$

مثال4.12: يدعي أمير صاحب مصنع المستقبل لإنتاج المصابيح الكهربية أن نسبة التالف في انتاجه أقل من نسبة التالف في مصنع الدقة لأحمد وبالتالي فهو يدعي أن انتاج مصنعه أفضل من انتاج مصنع أحمد. ولاختبار هذا الادعاء أخذت عينة من كل مصنع وفحصت العينتان وتم عد القطع التالفة في كل مصنع ، ويبين الجدول التالي نتائج هذه التجربة.

	مصنع المستقبل	مصنع الدقة
حجم العينة	50	100
عدد القطع التالفة	4	5

هل تدعم هذه البيانات صحة ادعاء امير عند مستوى معنوية 0.05

الحل/

نحسب نسب التالف في العينتين:

$$\hat{p}_1 = \frac{4}{50} = 0.08 \qquad (\text{مصنع المستقبل})$$

$$\hat{p}_2 = \frac{5}{100} = 0.05 \qquad (\text{مصنع الدقة})$$

<u>الفرضيات:</u>

$$H_0: P_1 - P_2 = 0$$
$$H_1: P_1 - P_2 > 0$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو التوزيع الطبيعي وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاه واحد نحو اليمين إذن تكون القيمة الحرجة:

$$Z_{\alpha} = 1.64$$

المعيار:

$$\mathrm{Z} = \frac{0.08 - 0.05}{\sqrt{\frac{0.08 \times 0.92}{50} + \frac{0.05 \times 0.95}{100}}} = 0.68$$

المقارنة والقرار:

الطريقة الاولى:

وحيث أن $(Z_{\alpha} = 1.64) < (Z = 0.68)$ فإننا نقبل الفرضية الصفرية وبالتالي لا يوجد دليل كافٍ على أن ادعاء امير صحيح .

الطريقة الثانية:

$$\mathrm{p-value} = P(Z > 0.68) = 0.2483 > 0.05$$

و بالتالي فإننا نقبل الفرضية الصفرية.

خامساً/ اختبار تباين مجتمع طبيعي التوزيع.

بفرض أن $X \sim \mathrm{N}(\mu, \sigma^2)$ و نريد اختبار ما إذا كان تباين المجتمع يساوي قيمة معينة σ_0^2 فيتم صياغة الفرضية الصفرية كالتالي:

$$H_0: \sigma^2 = \sigma_0^2$$

وتأخذ الفرضية البديلة إحدى الصور:

$$H_1: \sigma^2 \neq \sigma_0^2$$
$$H_1: \sigma^2 > \sigma_0^2$$
$$H_1: \sigma^2 < \sigma_0^2$$

ويكون المعيار المستخدم هو:

$$\mathbf{C} = \frac{(\boldsymbol{n}-\mathbf{1})\boldsymbol{S}^{\mathbf{2}}}{\boldsymbol{\sigma}_{\mathbf{0}}^{\mathbf{2}}} \sim \chi^{\mathbf{2}}_{(\boldsymbol{n}-\mathbf{1})}$$

مثال4.13: معمل لإنتاج الأنابيب البلاستيكية يسوق انتاجه من الأنابيب عندما يكون الانحراف المعياري في سمك جدار الأنبوب بحدود القيمة 0.003سم. قام مفتش وزارة الصناعة بسحب عينة عشوائية من انتاج يوم ما عبارة عن 20 انبوب فلوحظ أن الانحراف المعياري في سمك جدار الانبوب كان 0.004سم. هل سيسمح المفتش للمصنع بتسويق الانابيب لذلك اليوم. استخدم مستوى معنوية 0.05

الحل/

الفرضيات:

$$H_0: \sigma^2 = (0.003)^2$$
$$H_1: \sigma^2 \neq (0.003)^2$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع $\chi^2_{(19)}$ وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاهين إذن تكون القيم الحرجة:

$$\chi^2_{(0.975,19)} = 8.907 \quad ، \quad \chi^2_{(0.025,19)} = 32.852$$

المعيار:

$$C = \frac{(n-1)S^2}{\sigma_0^2} = \frac{19 \times (0.004)^2}{(0.003)^2} = 33.78$$

المقارنة والقرار:

الطريقة الاولى:

حيث أن $(C = 33.78) > (\chi^2_{(0.025,19)} = 32.852)$ فإننا نرفض الفرضية الصفرية وبالتالي المصنع مخالف للمواصفات ولن يسمح المفتش بتسويق انتاج ذلك اليوم .

الطريقة الثانية:

$$\text{p} - \text{value} = 2 \times P\left(\chi^2_{(19)} > 33.78\right) = 0.02 < 0.05$$

و بالتالي فإننا نرفض الفرضية الصفرية.

سادساً/ اختبار تساوي تبايني مجتمعين مستقلين طبيعيي التوزيع.

في هذه الحالة تكون الفرضية الصفرية على الصورة:

$$H_0: \sigma_1^2 = \sigma_2^2$$

ويتم وضعها على الصورة التالية والتي سيتم استخدامها لسهولة التعامل معها:

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$$

وتأخذ الفرضية البديلة إحدى الصور:

$$H_1: H_0: \frac{\sigma_1^2}{\sigma_2^2} \neq 1$$

$$H_1: H_0: \frac{\sigma_1^2}{\sigma_2^2} > 1$$

ويكون المعيار المستخدم في هاتين الحالتين هو:

$$\mathbf{F} = \frac{\boldsymbol{S_1^2}}{\boldsymbol{S_2^2}} \sim \mathbf{F}(\boldsymbol{n_1} - \mathbf{1}, \boldsymbol{n_2} - \mathbf{1})$$

$$H_1: H_0: \frac{\sigma_2^2}{\sigma_1^2} > 1$$

ويكون المعيار المستخدم في هذه الحالة هو:

$$\mathbf{F} = \frac{\boldsymbol{S_2^2}}{\boldsymbol{S_1^2}} \sim \mathbf{F}(\boldsymbol{n_2} - \mathbf{1}, \boldsymbol{n_1} - \mathbf{1})$$

مثال4.14: في دراسة لبيان معنوية الفرق بين تبايني أوزان الأرانب بعمر 6 أشهر في المزرعتين A و B تم سحب عينة عشوائية من كل مزرعة فكانت النتائج بالكيلوجرام كالتالي:

المزرعة A	1.5	1.7	1.3	1.4	1.1	1.4	1.2
المزرعة B	1.7	1.6	1.3	1.4	1.4	1.5	

اختبر تساوي تبايني الأوزان في المزرعتين من عدمه عند مستوى معنوية 0.05

الحل/ من الجدول نجد أن:

$$S_1^2 = 0.039 \quad , \quad S_2^2 = 0.022$$

الفرضيات:

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$$
$$H_1: \frac{\sigma_1^2}{\sigma_2^2} \neq 1$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع $F(6,5)$ وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاهين إذن تكون القيم الحرجة:

$$\frac{1}{F(0.025,5,6)} = \frac{1}{5.9876} = 0.167 \quad ، F(0.025,6,5) = 6.9777$$

المعيار:

$$F = \frac{S_1^2}{S_2^2} = \frac{0.039}{0.022} = 1.773$$

المقارنة والقرار:

الطريقة الاولى:

حيث أن $\quad 0.167 \leq (F = 1.773) \leq 6.9777$

فإننا نقبل الفرضية الصفرية وبالتالي لا فرق بين تباين اوزان المزرعتين .

الطريقة الثانية:

$$p - \text{value} = 2 \times P(F(6,5) < 1.773) = 0.2 > 0.05$$

و بالتالي فإننا نقبل الفرضية الصفرية.

مثال4.15: من مجتمع طبيعي التوزيع سحبت عينة عشوائية بحجم 15 مشاهدة بلغ تباينها 36 ، ومن مجتمع آخر طبيعي التوزيع سحبت عينة عشوائية بحجم 18 مشاهدة بلغ تباينها 25 .فهل تباين المجتمع الاول مساوٍ لتباين المجتمع الثاني أم أكبر منه. استخدم معنوية 10%

الحل/

من المسألة نجد أن $n_1 = 15 \quad , \quad \sigma_1^2 = 36 \quad , \quad n_2 = 18 \quad , \quad \sigma_2^2 = 25$

الفرضيات:

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$$

$$H_1: H_0: \frac{\sigma_1^2}{\sigma_2^2} > 1$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع $F(14,17)$ وحيث أن $\alpha = 0.1$ والاختبار ذو اتجاه واحد نحو اليمين تكون القيمة الحرجة:

$$F(0.1,14,17) = 1.91$$

المعيار:

$$F = \frac{S_1^2}{S_2^2} = \frac{36}{25} = 1.44$$

المقارنة والقرار:

الطريقة الأولى:

وحيث أن $F = 1.44 \leq F(0.1,14,17) = 1.91$

إذن نقبل الفرضية الصفرية وتباين المجتمعين متساوٍ.

الطريقة الثانية:

$$p - \text{value} = P(F(14,17) > 1.44) = 0.1 = \alpha$$

و بالتالي فإننا نقبل الفرضية الصفرية.

سابعاً/ اختبار بارتليت (Bartlett Test)

يستخدم هذا الاختبار لبحث تساوي تباينات عدة مجتمعات مستقلة عددها r عن طريق سحب عينة حجمها n_i حيث $i = 1, 2, 3, \cdots, r$ من كل مجتمع، ولا يشترط تساوي حجوم العينات المسحوبة، ويتبع هذا الاختبار توزيع χ^2 وهو اختبار من جهة واحدة نحو اليمين، وهو يختبر الفرضيات التالية:

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \cdots\cdots\cdots = \sigma_r^2$$

H_1: على الاقل أحد التباينات مختلف

والمعيار المستخدم في هذا الاختبار هو:

$$\mathbf{B} = \frac{(n-r)\ln\left[\frac{1}{n-r}\sum_{i=1}^{r}(n_i-1)S_i^2\right] - \sum_{i=1}^{r}(n_i-1)\ln S_i^2}{1+\frac{1}{3(r-1)\left[\sum_{i=1}^{r}\frac{1}{n_i-1}-\frac{1}{n-r}\right]}}$$

حيث أن:

$n = \sum_{i=1}^{r} n_i$: مجموع المشاهدات في كل العينات.

S_i^2: تباين العينة i

ويكون المعيار عبارة عن متغير عشوائي يتبع توزيع χ^2 بدرجة حرية $r-1$ بمعنى:

$$\mathbf{B} \sim \chi^2(r-1)$$

مثال4.16: في تجربة لبيان تأثير نوع الغذاء على زيادة الوزن عند الأبقار، تم سحب ثلاث عينات عشوائية من ثلاثة مزارع C, B, A . وبعد انتهاء مدة التجربة لوحظت الزيادات التالية في أوزان الابقار بالكغم وكانت كالتالي:

		4	7	6	6	A
5	1	3	5	3	4	B
	8	6	8	9	5	C

اختبر معنوية الفرق بين تباينات المجتمعات الثلاث بمستوى معنوية 0.05

الحل/

قبل البدء في الاختبار نكون الجدول التالي لحساب المعيار:

	A	B	C	المجموع
n_i	4	6	5	15
$\sum x_i$	23	21	36	
$\sum x_i^2$	137	85	270	
S_i^2	1.583	2.3	2.7	
$n_i - 1$	3	5	4	12
$(n_i - 1)S_i^2$	4.749	11.5	10.8	27.049
$\ln S_i^2$	0.459	0.833	0.993	
$(n_i - 1)\ln S_i^2$	1.377	4.465	3.972	9.514

<u>الفرضيات:</u>

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$$

على الاقل أحد التباينات مختلف: H_1

<u>التوزيع و القيم الحرجة المعيارية:</u>

التوزيع المستخدم هو توزيع $\chi^2_{(2)}$ وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاه واحد نحو اليمين إذن تكون القيمة الحرجة:

$$\chi^2_{(0.05,2)} = 5.991$$

<u>المعيار:</u>

$$B = \frac{(15-3)\ln\left[\frac{1}{15-3} \times 27.049\right] - 9.514}{1 + \frac{1}{3(3-1)\left[\frac{1}{3} + \frac{1}{5} + \frac{1}{4} - \frac{1}{15-3}\right]}} = 0.217$$

<u>المقارنة والقرار:</u>

الطريقة الاولى:

وحيث أن $(\chi^2_{(0.05,2)} = 5.991) > (B = 0.217)$ فإننا نقبل الفرضية الصفرية وبالتالي المجتمعات الثلاث لها نفس التباين .

الطريقة الثانية:

$$\text{p}-\text{value}=P\left(\chi^2_{(19)}>0.217\right)=0.995>0.05$$

إذن نقبل الفرضية الصفرية

ثامناً/ اختبار جودة المطابقة Goodness – of fit test

يسمى هذا الاختبار أيضاً باختبار جودة التوافق ويستخدم لبحث انتماء ظاهرة معينة لتوزيع احتمالي معين كأن نقول أن درجات الطلاب تتبع توزيعاً طبيعياً أو أن نقول أن عدد الوحدات المنتجة في خط انتاجي من معمل الالبان تتبع توزيع ذي الحدين وهكذا.

تتلخص فكرة جودة مطابقة توزيع متغير ما لتوزيع معين في مقارنة عدد المشاهدات التي نراها في التجربة مع عدد المشاهدات المتوقعة بفرض صحة الادعاء بأن العينة مسحوبة من مجتمع يتبع ذلك التوزيع المعين. وتم صياغة معيار يعبر عن الفرق بين التكرارات المشاهدة والتكرارات المتوقعة، والمعيار لهذا الاختبار في حالة التوزيعات المنفصلة هو:

$$\mathbf{C}=\sum_{i=1}^{k}\frac{(|O_i-E_i|-\frac{1}{2})^2}{E_i}\sim\chi^2_{(k-1)}$$

و في حالة التوزيعات المتصلة هو:

$$\mathbf{C}=\sum_{i=1}^{k}\frac{(O_i-E_i)^2}{E_i}\sim\chi^2_{(k-1)}$$

حيث:

k : تشير الى عدد مستويات العامل المدروس (عدد فئات جدول التوزيع التكراري للظاهرة)

O_i : التكرار المشاهد عند الفئة i

E_i : التكرار المتوقع عند الفئة i

والاختبار هو من اتجاه واحد نحو اليمين، وهو يختبر الفرضيات التالية:

H_0: **العينة المسحوبة تتبع توزيع** (يتم تحديد التوزيع المطلوب)

H_1: **العينة المسحوبة لا تتبع توزيع** (يتم تحديد التوزيع المطلوب)

والجدير بالذكر أنه يُوصى بتطبيق الاختبار إذا كان حجم العينة أكبر من 50 مشاهدة وأن لا يقل التكرار المشاهد في كل فئة عن 5 مشاهدات وفي حالة وجود فئة تحتوي على أقل من 5 مشاهدات تدمج مع الفئة المجاورة مع انقاص درجة الحرية بمقدار 1 عن كل فئة مدمجة.

مثال4.17: البيانات التالية تمثل عدد السيارات المستفيدة من خدمات محطة بنزين المدينة A ولستة أيام. تحقق من مدى مطابقة هذه المعطيات مع الفرضية القائلة بتساوي عدد السيارات لكل يوم المستفيدة من خدمات هذه المحطة عند مستوى معنوية 0.01

الايام	1	2	3	4	5	6	المجموع
عدد السيارات	120	90	90	110	120	130	660

الحل/

لاحظ أن الفئات هنا هي الايام وعددها 6 ومجموع التكرارات أكبر من 50 وكل فئة تكرارها أكبر من 5 و حيث أن التوزيع المطلوب اجراء الاختبار له هو التوزيع المنتظم وحيث أن $k = 6$ فيكون احتمال كل فترة متساوٍ ويساوي $\frac{1}{6}$ وحيث أن مجموع التكرارات 660 فيكون التكرار المتوقع لكل يوم هو:

$$E_i = \frac{\text{عدد السيارات}}{\text{عدد الايام}} = \frac{660}{6} = 110$$

جدول التوافق

O_i	E_i	$\lvert O_i - E_i \rvert - \frac{1}{2}$	$(\lvert O_i - E_i \rvert - \frac{1}{2})^2$	$\frac{(\lvert O_i - E_i \rvert - \frac{1}{2})^2}{E_i}$
120	110	9.5	90.25	0.8205
90	110	19.5	380.25	3.4568
90	110	19.5	380.25	3.4568
110	110	−0.5	0.25	0.0023
120	110	9.5	90.25	0.8205
130	110	19.5	380.25	3.4568
			المجموع	12.0137

الفرضيات:

$$H_0: X \sim \mathrm{U}(6)$$
$$H_1: X \ \downarrow\sim \ \mathrm{U}(6)$$

حيث نقصد بالرمز $\sim\downarrow$ عبارة "**لا تتبع**"

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع $\chi^2_{(5)}$ وحيث أن $\alpha = 0.01$ والاختبار ذو اتجاه واحد نحو اليمين إذن تكون القيمة الحرجة:

$$\chi^2_{(0.01,5)} = 15.086$$

المعيار:

حيث أن التوزيع المطلوب هو التوزيع المنتظم وهو من النوع المنفصل فيكون المعيار هو:

$$\mathrm{C} = \sum_{i=1}^{k} \frac{(|O_i - E_i| - \frac{1}{2})^2}{E_i} = 12.0137$$

المقارنة والقرار:

الطريقة الاولى:

وحيث أن $\mathrm{C} = 12.0137 < (\chi^2_{(0.01,5)} = 15.086)$ فإننا نقبل الفرضية الصفرية وبالتالي توزيع السيارات هو توزيع منتظم .

الطريقة الثانية:

$$\mathrm{p-value} = P\left(\chi^2_{(5)} > 12.0137\right) = 0.025 > 0.01$$

إذن نقبل الفرضية الصفرية

مثال4.18: هل العينة العشوائية التالية تشير إلى أن المكالمات الهاتفية القادمة من أستراليا هي بتوزيع بواسون وذلك عند مستوى معنوية 0.05

عدد المكالمات	0	1	2	3	4	5
عدد الايام	100	20	15	6	6	3

الحل/

لاحظ أن الفئات هنا هي عدد المكالمات وعددها 6 ومجموع التكرارات أكبر من 50 وكل فئة تكرارها أكبر من 5 عدا الفئة السادسة فتكرارها 3 فيتم دمجها مع الفئة الخامسة فيصبح الجدول كالتالي:

عدد المكالمات x_i	0	1	2	3	4	$n=\sum f_i$
عدد الايام f_i	100	20	15	6	9	150

حيث أن التوزيع المطلوب اجراء الاختبار له هو توزيع بواسون وهو من النوع المنفصل والاقتران الاحتمالي له هو:

$$P(x)=\frac{e^{-\lambda}(\lambda)^x}{x!}$$

وحيث أن معلمة التوزيع مجهولة نقدرها كالتالي:

$$\hat{\lambda}=\frac{\sum x_i f_i}{\sum f_i}=\frac{100\times 0+20\times 1+15\times 2+6\times 3+9\times 4}{150}=0.69$$

فيكون الاقتران الاحتمالي لتوزيع بواسون هو $P(x)=\frac{e^{-0.69}(0.69)^x}{x!}$

وتكون الاحتمالات المناظرة للفئات كما هو موضح بالجدول التالي:

عدد المكالمات x_i	0	1	2	3	4
عدد الايام f_i	100	20	15	6	9
$P_r(X=x_i)$	0.5	0.35	0.119	0.027	0.004

جدول التوافق

O_i	$E_i=nP_r(X=x_i)$	$\lvert O_i-E_i\rvert-\frac{1}{2}$	$(\lvert O_i-E_i\rvert-\frac{1}{2})^2$	$\frac{(\lvert O_i-E_i\rvert-\frac{1}{2})^2}{E_i}$
100	75	24.5	600.25	8.0033
20	52.5	32	1024	19.5048
15	17.85	2.35	5.5225	0.3093
6	4.05	1.45	2.1025	0.5191
9	0.6	7.9	62.41	104.0167
			المجموع	132.3532

الفرضيات:

$$H_0 : X \sim \text{PO}(0.69)$$
$$H_1 : X \ \not\sim \ \text{PO}(0.69)$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع $\chi^2_{(4)}$ ولاحظ أننا دمجنا فئة وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاه واحد نحو اليمين إذن تكون القيمة الحرجة:

$$\chi^2_{(0.05,4)} = 9.488$$

المعيار:

$$\text{C} = \sum_{i=1}^{k} \frac{(|O_i - E_i| - \frac{1}{2})^2}{E_i} = 132.3532$$

المقارنة والقرار:

الطريقة الاولى:

وحيث أن $(\chi^2_{(0.05,4)} = 9.844) > \text{C} = 132.3532$ فإننا نرفض الفرضية الصفرية وبالتالي مجتمع الدراسة لا يتبع توزيع بواسون .

الطريقة الثانية:

$$\text{p} - \text{value} = P\left(\chi^2_{(4)} > 132.3532\right) = 0.005 < 0.05$$

إذن نرفض الفرضية الصفرية

مثال4.19: لديك التوزيع التكراري التالي، بين عند مستوى معنوية 0.05 إذا ما كانت البيانات التي في الجدول هي من مجتمع طبيعي التوزيع.

الفئات	2−	4−	6−	8−	12−10
التكرار	6	22	30	20	8

الحل/

قبل اجراء الاختبار سنحسب الوسط الحسابي والتباين للتوزيع التكراري .

مركز الفئة x_i	التكرار f_i	$x_i f_i$	$(x_i-\overline{x})^2 f_i$
3	6	18	98.415
5	22	110	92.455
7	30	210	0.075
9	20	180	76.05
11	8	88	124.82
المجموع	$n = 86$	606	391.815

$$\hat{\mu} = \frac{\sum x_i f_i}{\sum f_i} = \frac{606}{86} = 7.05$$

$$\hat{\sigma} = S = \sqrt{\frac{\sum(x_i-\overline{x})^2 f_i}{\sum f_i - 1}} = \sqrt{\frac{391.815}{86-1}} = 2.147$$

نحول الحدود الفعلية للفئات للدرجة المعيارية باستخدام العلاقة $z_i = \frac{UB-\hat{\mu}}{\hat{\sigma}}$ ثم نحسب احتمال كل فئة كما هو موضح بالجدول التالي:

الفئة	$f(z_i)$	$P(z < z_i)$	z_i	الحدود الفعلية للفئات UB
−	−	0.0094	−2.35	2
−2	0.0684	0.0778	−1.42	4
−4	0.2343	0.3121	−0.49	6
−6	0.3579	0.67	0.44	8
−8	0.2447	0.9147	1.37	10
10−12	0.0749	0.99	2.30	12

حيث احتمال الفئة يحسب من العلاقة $f(z_i) = P(z < z_i) - P(z < z_{i-1})$

جدول التوافق

O_i	$E_i = nf(z_i)$	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
6	5.88	0.12	0.0144	0.0024
22	20.15	1.85	3.4225	0.1699
30	30.78	−0.78	0.6084	0.0198
20	19.32	0.68	0.4624	0.0239
8	6.44	1.56	2.4336	0.3779
			المجموع	0.5939

الفرضيات:

$$H_0{:}X \sim \mathrm{N}(7.05{,}2.147)$$
$$H_1{:}X \downarrow\sim \mathrm{N}(7.05{,}2.147)$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع $\chi^2_{(4)}$ وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاه واحد نحو اليمين إذن تكون القيمة الحرجة:

$$\chi^2_{(0.05,4)} = 9.488$$

المعيار:

$$\mathrm{C} = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = 0.5939$$

المقارنة والقرار:

الطريقة الاولى:

وحيث أن $(\chi^2_{(0.05,4)} = 9.844 > \mathrm{C} = 0.5939$ فإننا نقبل الفرضية الصفرية وبالتالي مجتمع الدراسة يمكن أن يتبع التوزيع الطبيعي .

الطريقة الثانية:

$$\mathrm{p-value} = P\left(\chi^2_{(4)} > 0.5939\right) = 0.975 > 0.05$$

إذن نقبل الفرضية الصفرية

مثال4.20: يمثل الجدول التالي التوزيع التكراري لعلامات 200 طالب في مادة التفاضل و التكامل في جامعة ما.

الفئات	التكرار
< 50	20
[50,59]	50
[60,69]	30
[70,79]	45
[80,99]	35
≥ 99	20
المجموع	$n = 200$

اختبر الفرضية المبدئية التي تقول أن البيانات المعطاة تتبع توزيعا طبيعيا وسطه 70 وانحراف معياري 10 وذلك على مستوى دلالة 0.05

الحل/

نكون الجدول التالي:

الحدود الفعلية للفئات b_i	$z_i = \frac{b_i - 70}{10}$	$P(z < z_i)$	$f(z_i)$
49.5	−2.05	0.0202	0.0202
59.5	−1.05	0.1469	0.1267
69.5	−0.05	0.5199	0.373
79.5	0.95	0.8289	0.309
89.5	1.95	0.9441	0.1152
	−1.95	0.0256	0.0256

جدول التوافق

O_i	$E_i = nf(z_i)$	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
20	4.04	15.9256	254.7216	63.0499
50	25.34	24.66	608.1156	23.9982
30	74.6	−44.6	1989.16	26.6643
45	61.8	−16.8	282.24	4.567
35	23.04	11.96	143.0416	6.2084
20	5.12	14.88	221.4144	43.245
			المجموع	167.7328

الفرضيات:

$$H_0{:}X \sim \mathrm{N}(70,100)$$
$$H_1{:}X \downarrow\sim \ \mathrm{N}(70,100)$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع $\chi^2_{(5)}$ وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاه واحد نحو اليمين إذن تكون القيمة الحرجة:

$$\chi^2_{(0.05,5)} = 11.07$$

المعيار:

$$\mathrm{C} = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = 167.7328$$

المقارنة والقرار:

الطريقة الاولى:

حيث أن $(\chi^2_{(0.05,5)} = 11.07) < \mathrm{C} = 167.7328$ فإننا نرفض الفرضية الصفرية وبالتالي مجتمع الدراسة لا يمكن أن يتبع التوزيع الطبيعي بوسط 70 وانحراف معياري 10 .

الطريقة الثانية:

$$\text{p}-\text{value}=P\left(\chi^2_{(5)}>167.7328\right)=0.005<0.05$$

إذن نرفض الفرضية الصفرية

تاسعاً/ الاختبارات الخاصة بتحليل الارتباط و الانحدار

1- اختبار معامل الانحدار الخطي البسيط

نعلم من مساق مبادئ الاحصاء كيفية ايجاد معادلة خط الانحدار بين متغيرين كميين x و y عن طريق سحب عينة عشوائية ثنائية القيم (x,y) وبحجم n ، حيث كانت معادلة خط الانحدار تحسب من العلاقة:

$$\boldsymbol{y=\vartheta x+\beta+\epsilon}$$

حيث يكون X هو المتغير المستقل و Y هو المتغير التابع و وتسمى $\boldsymbol{\vartheta}$ بميل خط انحدار Y على X وتكون $\boldsymbol{\beta}$ هي مقطع خط الانحدار على محور Y أما $\epsilon\sim N(0,\sigma^2)$ فتعبر عن الخطأ في التقدير.

تلعب $\boldsymbol{\vartheta}$ دورا مهماً في التحليل الاحصائي، حيث أنها تعبر عن معدل التغير في Y بالنسبة لـ X . فإذا كانت X تمثل مصروفات الدعاية على نوع من السيارات وكانت Y تمثل قيمة المبيعات الشهرية من السيارات، فتكون $\boldsymbol{\vartheta}$ هي معدل الزيادة في المبيعات بالنسبة لمصروفات الدعاية.

ويتم تقدير $\boldsymbol{\vartheta}$ من العلاقة:

$$\boldsymbol{\hat{\vartheta}=\frac{SS_{xy}}{SS_x}}$$

حيث:

$$\boldsymbol{SS_{xy}=\sum_{i=1}^{n}x_iy_i-n\bar{x}\,\bar{y}}$$

$$\boldsymbol{SS_x=\sum_{i=1}^{n}x_i^2-n\bar{x}^2}$$

والجدير بالذكر أنه سيعتمد اختبار الفرضيات حول $\boldsymbol{\vartheta}$ على $\boldsymbol{\hat{\vartheta}}$ للقيام بعملية الاختبار، حيث يمكننا مما سبق استنتاج أن توزيع $\boldsymbol{\hat{\vartheta}}$ هو توزيع طبيعي وأن:

$$\mathbf{E}\boldsymbol{(\hat{\vartheta})=\vartheta}$$

$$\boldsymbol{var(\hat{\vartheta})=\sigma^2_{\hat{\vartheta}}=\frac{\sigma^2}{SS_x}}$$

حيث يسمى $\sigma_{\hat{\vartheta}}$ الخطأ المعياري للمتغير العشوائي $\hat{\vartheta}$

و في حالة مجهولية σ^2 يقدر تباين المجتمع بتباين العينة المسحوبة منه S حيث يكون:

$$\hat{\sigma}^2_{\hat{\vartheta}} = \frac{S^2}{SS_x}$$

في بعض المسائل العملية يكون هدفنا معرفة إذا ما كانت Y تزداد أو تنقص خطياً بزيادة أو بنقصان X أم لا، بمعنى هل هناك علاقة خطية غير ثابتة بين المتغيرين أم لا، وهذا التساؤل يمكننا من صياغة الفرضيات التالية للدلالة على هذه التساؤلات:

$$H_0: \vartheta = \vartheta_0$$
$$H_1: \vartheta \neq \vartheta_0$$
$$H_1: \vartheta > \vartheta_0$$
$$H_1: \vartheta < \vartheta_0$$

إن المعيار المستخدم في اختبار الفرضيات السابقة هو:

$$\mathbf{T} = \frac{\hat{\vartheta} - \vartheta_0}{S / \sqrt{SS_x}} \sim t(n-2)$$

حيث:

$$S^2 = \frac{SS_E}{n-2}$$
$$SS_E = SS_y - \hat{\vartheta}\, SS_{xy}$$
$$SS_y = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2$$

مثال4.21: للبيانات المعطاة في الجدول التالي، اختبر ما إذا كانت $\vartheta = 0$ أم لا على مستوى دلالة 0.05

x	43	21	64	57	47	28	75	34	52	39
y	78	52	85	92	89	73	95	56	75	65

الحل/

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 46$$

$$\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = 76$$

$$\sum_{i=1}^{n} x_i^2 = 23634$$

$$\sum_{i=1}^{n} y_i^2 = 59378$$

$$\sum_{i=1}^{n} x_i y_i = 36821$$

وبالتالي:

$$SS_x = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2 = 2474$$

$$SS_y = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 = 1618$$

$$SS_{xy} = \sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y} = 1861$$

$$\hat{\vartheta} = \frac{SS_{xy}}{SS_x} = \frac{1861}{2474} = 0.72$$

$$SS_E = SS_y - \widehat{\vartheta}\, SS_{xy} = 1618 - 0.72 \times 1861 = 278.08$$

$$S^2 = \frac{SS_E}{n-2} = \frac{278.08}{8} = 34.76$$

<u>الفرضيات:</u>

$$H_0: \vartheta = 0$$
$$H_1: \vartheta \neq 0$$

<u>التوزيع و القيم الحرجة المعيارية:</u>

التوزيع المستخدم هو توزيع $t(8)$ وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاهين إذن تكون القيم الحرجة:

$$-t(0.025,8) = -2.306 \quad ، \quad t(0.025,8) = 2.306$$

المعيار:

$$T = \frac{\hat{\vartheta} - \vartheta_0}{S/\sqrt{SS_x}} = \frac{0.72}{5.9/\sqrt{2474}} = 6.07$$

المقارنة والقرار:

الطريقة الاولى:

حيث أن $T = 6.07 > (t(0.025,8) = 2.306)$ فإننا نرفض الفرضية الصفرية وبالتالي هناك دليل كافٍ على أن $\vartheta \neq 0$.

الطريقة الثانية:

$$\text{p} - \text{value} = 2 \times P(t > 3.78) = 0.01 < 0.05$$

و بالتالي فإننا نرفض الفرضية الصفرية.

2- اختبار مقطع خط الانحدار الخطي البسيط

يُعرف مقطع خط الانحدار β بأنه متوسط المتغير y عندما تكون $x = 0$.

إن توقع المتغير Y عند قيمة معينة للمتغير X ولتكن $\boldsymbol{x_0}$ يعطى بالعلاقة:

$$E(Y/x_0) = \vartheta x_0 + \beta$$

ويقدر هذا التوقع بالمقدار

$$E(\hat{Y}/x_0) = \hat{\vartheta} x_0 + \hat{\beta}$$

ويكون

$$E(\hat{Y}/x_0) \sim \mathrm{N}\left(\vartheta x_0 + \beta, \sigma^2 \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SS_x}\right]\right)$$

وبالتالي:

$$\hat{\beta} \sim \mathrm{N}\left(\beta, \sigma^2 \left[\frac{1}{n} + \frac{(\bar{x})^2}{SS_x}\right]\right)$$

و في حالة مجهولية σ^2 يقدر تباين المجتمع بتباين العينة المسحوبة منه S وتحسب $\hat{\beta}$ من العلاقة:

$$\boldsymbol{\hat{\beta} = \bar{y} - \hat{\vartheta}\bar{x}}$$

تكون الفرضيات في هذا الاختبار كالتالي:

$$H_0: \beta = \beta_0$$
$$H_1: \beta \neq \beta_0$$
$$H_1: \beta > \beta_0$$
$$H_1: \beta < \beta_0$$

ويكون المعيار

$$T = \frac{\hat{\beta} - \beta_0}{\sqrt{S^2 \left[\frac{1}{n} + \frac{(\bar{x})^2}{SS_x}\right]}} \sim t(n-2)$$

مثال4.22: في دراسة للعلاقة بين كمية الأمطار التي نزلت وارتفاع شجرة الزيتون بعمر معين، سحبت عينة عشوائية من 6 مناطق فأعطت النتائج التالية:

كمية الامطار X	5	10	15	20	25	30
متوسط الطول بالمتر Y	3	3	4	6	6	7

اختبر معنوية مقطع الانحدار عند مستوى معنوية 0.05

الحل/

	x	y	xy	x^2	y^2
	5	3	15	25	9
	10	3	30	100	9
	15	4	60	225	16
	20	6	120	400	36
	25	6	150	625	36
	30	7	210	900	49
المجموع	105	29	585	2275	155

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 17.5$$

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{29}{6}$$

$$SS_x = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2 = 437.5$$

$$SS_y = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 = 14.83$$

$$SS_{xy} = \sum_{i=1}^{n} x_i y_i - n\bar{x}\,\bar{y} = 77.5$$

$$\hat{\vartheta} = \frac{SS_{xy}}{SS_x} = \frac{77.5}{437.5} = 0.177$$

$$\hat{\beta} = \bar{y} - \hat{\vartheta}\bar{x} = \frac{29}{6} - 0.177 \times 17.5 = 1.736$$

$$SS_E = SS_y - \widehat{\vartheta}\, SS_{xy} = 14.83 - 0.177 \times 77.5 = 1.1125$$

$$S^2 = \frac{SS_E}{n-2} = \frac{1.1125}{4} = 0.2781$$

الفرضيات:

$$H_0 : \beta = 0$$
$$H_1 : \beta \neq 0$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع $t(4)$ وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاهين إذن تكون القيم الحرجة:

$$-\,t(0.025{,}4) = -2.78 \quad ، \quad t(0.025{,}4) = 2.78$$

المعيار:

$$\mathrm{T} = \frac{\hat{\beta}}{\sqrt{S^2\left[\frac{1}{n} + \frac{(\bar{x})^2}{SS_x}\right]}} = \frac{1.737}{\sqrt{0.2781\left[\frac{1}{6} + \frac{(17.5)^2}{437.5}\right]}} = 3.538$$

المقارنة والقرار:

الطريقة الاولى:

حيث أن $\mathrm{T} = 3.538 > (\,t(0.025{,}4) = 2.78)$ فإننا نرفض الفرضية الصفرية وبالتالي هناك دليل كافٍ على أن $\beta \neq 0$.

الطريقة الثانية:

$$\mathrm{p-value} = 2 \times P(t > 3.538) = 0.02 < 0.05$$

و بالتالي فإننا نرفض الفرضية الصفرية.

3-اختبار معامل الارتباط الخطي البسيط

معامل ارتباط بيرسون $(\boldsymbol{\rho})$ بين متغيرين كميين x و y هو أحد مقاييس الارتباط الخطي بين متغيرين وتقع قيمته دائماً في الفترة $[-1,1]$ ويحسب من العلاقة:

$$\boldsymbol{\rho} = \frac{\mathbf{E}(\boldsymbol{XY}) - \mathbf{E}(\boldsymbol{X})\mathbf{E}(\boldsymbol{Y})}{\boldsymbol{\sigma_x \sigma_y}}$$

ويتم تقدير قيمته بسحب عينة عشوائية ثنائية القيم (x, y) وبحجم n ويحسب من العلاقة:

$$r = \frac{SS_{xy}}{\sqrt{SS_x SS_y}} = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y}}{\sqrt{(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2)(\sum_{i=1}^{n} y_i^2 - n\overline{y}^2)}}$$

لاحظ أن هناك علاقة بين **r** وبين ميل خط الانحدار $\widehat{\boldsymbol{\vartheta}}$ وهذه العلاقة هي:

$$r = \sqrt{\frac{SS_x}{SS_y}}\ \widehat{\boldsymbol{\vartheta}}$$

وهذه العلاقة مكافئة للعلاقة

$$r^2 = \frac{SS_y - SS_E}{SS_y}$$

فإذا أردنا تقدير قيمة مستقبلة للمتغير y بناءً على عينة عشوائية $\{\boldsymbol{y_i}\}_{\boldsymbol{i=1}}^{\boldsymbol{n}}$ فمن الطبيعي أن نستخدم المقدر $\bar{\mathrm{y}}$ ويقاس الخطأ في التقدير في هذه الحالة بالمقدار:

$$SS_y = \sum (\boldsymbol{y_i} - \bar{\mathrm{y}})^2$$

ولكن عند وجود متغير آخر x مرتبط بالمتغير y وعند توفر معلومات عن المتغير x فإننا نستخدم معادلة خط الانحدار $\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{\vartheta}}\boldsymbol{x} + \widehat{\boldsymbol{\beta}}$ لتقدير قيمة y إذا علمت قيمة x ويكون الخطأ في التقدير

$$SS_E = \sum (\boldsymbol{y_i} - \hat{\mathrm{y}})^2$$

وبالتالي فإن العلاقة $r^2 = \frac{SS_y - SS_E}{SS_y}$ تعبر عن نسبة انخفاض الخطأ في التقدير نتيجة لاستخدام معادلة خط الانحدار وبالتالي فهي تعطي مؤشراً لمدى فاعلية استخدام معادلة خط الانحدار للتنبؤ بقيم المتغير y بدلاً من $\bar{y}$.

فمثلاً إذا كانت قيمة معامل الارتباط $r = 0.9$ فإن $r^2 = 0.81$ أي أن استخدام معادلة خط الانحدار في التقدير يقلل الخطأ في التقدير بنسبة 81% ، ومن الواضح أنه كلما كان الارتباط بين المتغيرين أقوى كلما كان استخدام معادلة خط الانحدار أفضل والخطأ في التقدير أقل، فلو كان الارتباط تاماً فإن الخطأ في التقدير يساوي صفر.

لهذا فإن قيمة $\boldsymbol{r^2}$ تحدد فعالية خط الانحدار في التقدير وعليه سميت $\boldsymbol{r^2}$ بمعامل التحديد Coefficient of Determination ، هذه يعطي سبباً للاهتمام بمعامل الارتباط، بالإضافة إلى أن حساب معامل الارتباط يستخدم في اختبار استقلالية المتغيرات.

فإذا أردنا أن نختبر أن المتغيران $\boldsymbol{x}$ و $\boldsymbol{y}$ مستقلان فإن الفرضيات المستخدمة هي:

$$H_0: \rho = 0$$
$$H_1: \rho \neq 0$$
$$H_1: \rho > 0$$
$$H_1: \rho < 0$$

والمعيار في هذه الحالة هو:

$$\mathbf{T} = \frac{\boldsymbol{r\sqrt{n-2}}}{\boldsymbol{\sqrt{1-r^2}}} \sim \boldsymbol{t(n-2)}$$

مثال4.23: اجريت دراسة على عشرين شاباً حيث تم رصد عُمر الشاب x وضغط دمه y وكانت النتائج كما يلي:

$$\sum_{i=1}^{n} x_i y_i = 113156 \quad , \quad \sum_{i=1}^{n} y_i^2 = 368946 \quad , \quad \sum_{i=1}^{n} x_i^2 = 36429$$

$$\sum_{i=1}^{n} x_i = 833 \quad , \quad \sum_{i=1}^{n} y_i = 2710 \quad , \quad n = 20$$

اختبر عند مستوى معنوية 0.05 وجود ارتباط بين عُمر الشاب وضغط دمه.

الحل/

$$r = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\,\bar{y}}{\sqrt{\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)\left(\sum_{i=1}^{n} y_i^2 - n\bar{y}^2\right)}}$$

$$= \frac{113156 - 20 \times \frac{833}{20} \times \frac{2710}{20}}{\sqrt{(36429 - 20\left(\frac{833}{20}\right)^2)(368946 - 20\left(\frac{2710}{20}\right)^2)}} = 0.1637$$

<u>الفرضيات:</u>

$$H_0{:}\,\rho = 0$$
$$H_1{:}\rho \neq 0$$

<u>التوزيع و القيم الحرجة المعيارية:</u>

التوزيع المستخدم هو توزيع $t(18)$ وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاهين إذن تكون القيم الحرجة:

$$-\,t(0.025{,}18) = -1.734 \quad ، \quad t(0.025{,}18) = 1.734$$

<u>المعيار:</u>

$$\mathrm{T} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.1637\sqrt{18}}{\sqrt{1-(0.1637)^2}} = 0.704$$

<u>المقارنة والقرار:</u>

الطريقة الاولى:

حيث أن $\mathrm{T} = 0.704 < (\, t(0.025{,}18) = 1.734)$ فإننا نقبل الفرضية الصفرية وبالتالي ليس هناك دليل كافٍ على وجود ارتباط ايجابي بين عمر الشاب وضغط دمه عند مستوى الدلالة المذكور.

الطريقة الثانية:

$$\mathrm{p-value} = 2 \times P(t > 0.704) = 0.2 > 0.05$$

و بالتالي فإننا نقبل الفرضية الصفرية.

تمارين:

(اتخذ القرار باستخدام المعنوية المحسوبة بعد حل كل سؤال)

1- أخذت عينة عشوائية حجمها 10 بوسط حسابي 812 من مجتمع طبيعي تباينه 150 اختبر عند مستوى معنوية 0.01 الفرضية $H_0: \mu = 800$ مقابل $H_1: \mu > 800$

2- يبلغ معدل طول الجندي في أحد الجيوش 169 سم وفي السنوات الاخيرة بدأ الاقبال على الجندية يزيد وبالتالي صارت القيادة تضع شروطاً أشد بخصوص القبول. اختبر الفرضية التي تقول أن معدل طول الجندي قد ازداد علماً بأن عينة عشوائية حجمها 50 من أفراد ذلك الجيش قد دُرِست وكان معدل الاطوال فيها 171.5 سم بانحراف معياري 5 وذلك عند مستوى معنوية 0.05 .

3- اظهرت دراسات سابقة أن التباين في أعمار المصابيح الكهربائية التي ينتجها أحد المصانع يساوي 170 ساعة تربيع. اختبر الفرضية $H_0: \mu = 850$ مقابل $H_1: \mu \neq 850$ إذا تم سحب عينة عشوائية من المصابيح أعطت وسطاً 839 ساعة وذلك عند مستوى معنوية 0.05 . احسب قوة الاختبار عندما $\mu = 800$

4- عينة عشوائية مكونة من 200 حبة بطاطا من بينها 80 حبة صغيرة و عينة عشوائية اخرى سحبت من مزرعة اخرى بحجم 300 حبة وجد بينها 100 حبة صغيرة. اختبر الفرضية فهل نسبة البطاطا الصغيرة في المزرعة الاولي أكبر منه في المزرعة الثانية . استخدم مستوى دلالة 0.05

5- معمل ينتج علب غذائية بوزن 500 غم حسب مقاييس الجودة، تم سحب عينة عشوائية مؤلفة من 15 علبة فكان الوسط الحسابي لأوزان العلب في العينة 488 غم بانحراف معياري 6 غم، اختبر عند مستوى معنوية 0.05 إذا ما كان المصنع ملتزم بمقاييس الجودة أم لا.

6- اعتاد رب اسرة أن يصرف ما معدله 4 دينار يومياً بانحراف معياري 1 دينار ولكنه لاحظ في الفترة الاخيرة زيادة المصروفات ولهذا الغرض قام برصد مصروفات 10 ايام وكانت كالتالي:

4 5 5 5 5 3 3 6 بالدينار فهل ما رصده يؤيد شكوكه عند مستوى معنوية 0.1

7- إذا كانت نسبة العائلات التي تملك البيوت التي تسكن فيها في مدينة معينة هي 62% ، اجريت دراسة على 2000 موظف فوجد أن 1280 شخص من بينهم يملكون البيوت التي يسكنوها. اختبر الفرضية $H_0: \mathrm{P} = 0.62$ مقابل $H_1: \mathrm{P} > 0.62$ عند معنوية 1%

8- إذا كانت علامات الطلبة في مادة الاحصاء تتبع توزيعاً طبيعياً وسطه μ وتباينه σ^2 واخذت عينة حجمها 20 طالباً من طلاب هذه المادة و اجري لهم اختبار فكان الوسط الحسابي لدرجاتهم 75 بانحراف معياري 9 أوجد فترة 90% ثقة للانحراف المعياري.

9- لسؤال 8 اختبر الفرضية $H_0: \sigma^2 = 82$ مقابل $H_1: \sigma^2 < 82$ عند نفس مستوى المعنوية، ثم احسب قوة الاختبار عندما $\sigma^2 = 78$ ثم عندما $\sigma^2 = 158$ فسر النتائج التي حصلت عليها.

10- يمثل الجدول التالي عدد الراسبين من الذكور والاناث في مبحث الفيزياء في امتحان الثانوية العامة في عينتين اخذتا من الطلبة الذين تقدموا للامتحان في العام 1985.

	اناث	ذكور
حجم العينة	40	50
عدد الراسبين	10	8

هل تستطيع أن تستنتج وجود فرق ذا دلالة احصائية على مستوى معنوية 0.05 بين نسبة النجاح في المجموعتين وإذا وجد فلصالح من.

11- اجريت دراسة لمقارنة دقة نوعين من الاجهزة في قياس مقدار التلوث في المياه العادمة لمصنع معين. وفي أحد الايام تم أخذ 7 قياسات بجهاز من نوع A و 6 قياسات من نوع B وكانت النتائج كما يلي:

A	0.99	0.86	0.71	0.96	0.78	0.82	0.95
B	0.89	0.9	0.94	0.94	0.91	0.89	

عند مستوى معنوية 0.05 حدد إذا ما كان احد الجهازين أدق من الآخر.

12-البيانات التالية تمثل كمية الغذاء المستهلك في اليوم الواحد من قبل الارانب وفي أشهر السنة المختلفة. اختبر الفرضية الصفرية القائلة بتجانس تباينات كمية الغذاء المستهلك على مدار اشهر السنة المختلفة عند مستوى معنوية 0.05

الشهر الثاني	الشهر الخامس	الشهر السابع	الشهر العاشر
4.2	4.3	4.5	4.6
5.2	4.1	4.3	5
4.3	4	4.2	5.1
4.7	4.1	4	4.8
4.2	3.8	4.3	5.3
	3.9	4.4	

13- استخدم المعلومات التالية لاختبار الفرضية $H_0{:}\,\vartheta = 0$ مقابل $H_1{:}\,\vartheta \neq 0$ عند مستوى دلالة 0.1

$$\sum_{i=1}^{n} x_i y_i = 12.48 \quad , \quad \sum_{i=1}^{n} y_i^2 = 11.8186 \quad , \quad \sum_{i=1}^{n} x_i^2 = 146$$

$$\sum_{i=1}^{n} x_i = 6 \quad , \quad \sum_{i=1}^{n} y_i = 11.74 \quad , \quad n = 12$$

14-يلخص الجدول التالي علامات 10 طلاب في امتحان اللغة الانجليزية حيث تم رصد ساعات الدراسة على الامتحان لكل طالب وعلامته في الامتحان.

ساعات الدراسة x	17	1	22	12	7	4	14	10	9	4
علامة الامتحان y	84	21	91	60	44	37	73	65	58	31

أ) اختبر صحة ادعاء أن معدل زيادة علامة الطالب لكل ساعة دراسة يساوي 4 علامات على الاقل عند معنوية 0.05

ب) هل تستطيع أن تستنتج أن مقطع خط الانحدار لا يساوي 30 عند نفس المعنوية السابقة.

15-في دراسة لبيان العلاقة بين انفاق الاسرة والدخل الشهري تم سحب عينة عشوائية حجمها 25 مشاهدة مزدوجة فظهر أن معامل الارتباط بينهما يساوي 0.32 ، اختبر معنوية معامل الارتباط عند مستوى 0.1

الفصل الخامس

تحليل التباين

Analysis of Variance (ANOVA)

درسنا في الفصل السابق اختبارات الفرضيات التي تقارن بين متوسطي مجتمعين، فإذا أردنا المقارنة بين متوسطات ثلاث مجتمعات علينا اجراء الاختبار ثلاث مرات وإذا أردنا المقارنة بين خمس مجتمعات علينا اجراء الاختبار 10 مرات وصفة عامة إذا أردنا المقارنة بين أوساط مجتمعات عددها k فهذا يتطلب اجراء الاختبار $\binom{k}{2}$ مرة ومن هنا تبرز مشكلتين، الأولى في كِبَرْ عدد مرات اجراء الاختبار والثانية وهي الأخطر وتتمثل في ضعف احتمال الوصول لقرار صحيح، حيث أنه بفرض أن مستوى المعنوية كان 5% – بمعنى أن احتمال الوصول لقرار صحيح هو 95% – وأردنا اجراء الاختبار على 5 مجتمعات فسيكون احتمال الوصول لقرار صحيح في العشر اختبارات مساوياً $0.599 = (0.95)^{10}$ أي 59.9% ومن هنا برزت الحاجة لأسلوب جديد لمقارنة متوسطات عدة مجتمعات وهذا الأسلوب يعرف باسم تحليل التباين.

وتكمن فكرة تحليل التباين في تجزئة التباين الكلي بين الاستجابات كلها (المشاهدات) في العينات جميعها إلى مكوناتها وذلك نسبة الى المصدر المسبب لهذا التباين، **وننوه هنا إلى أن هذا الاسلوب يختبر فقط وجود اختلاف بين أحد أوساط المجتمعات أو بعض الاوساط دون تحديد المجتمعات التي بينها اختلافات.**

لتطبيق هذا الأسلوب في التحليل يجب أن تتحقق بعض الشروط وهي:

1− المجتمعات قيد الدراسة تتبع التوزيع الطبيعي أي: $X_i \sim N(\mu_i, \sigma_i^2)$

2− جميع المجتمعات مستقلة عن بعضها البعض.

3− تساوي تباينات جميع المجتمعات، وبالتالي يمكن كتابة $X_i \sim N(\mu_i, \sigma^2)$

وفي البداية سنذكر بعض المصطلحات التي سنستخدمها في دراستنا للموضوع:

المادة التجريبية Experimental Material: وهي الشى الذي سنستخدمه في اجراء التجربة كأن تكون أرض واسعة أو مزرعة أبقار كبيرة.

الوحدة التجريبية Experimental Unit: وهي جزء من المادة التجريبية الذي سنستخدمه في اجراء التجربة كأن تكون قطعة أرض نريد أن نزرعها بالقمح لإجراء تجارب على القمح، أو بعض الأبقار نريد اجراء تجارب عليها لدراسة تأثير نوع معين من الأعلاف.

المعالجة Treatment : هي الطريقة (المجتمع) التي نقيس تأثيرها على الوحدة التجريبية، وقد تكون وصفية مثل مجموعة أصناف من القمح أو مجموعة من الأسمدة وقد تكون كمية كأن تكون عبارة عن مستويات لعامل **(Factor)** واحد كأن يكون مستويات تركيز لمبيد حشري حيث يكون المبيد هو العامل و المستويات هي المعالجات.

وحدة المعاينة Sampling Unit : وهي الجزء من الوحدة التجريبية الذي يُؤْخَذ لقياس تأثير المعالجة عليه، وقد تكون وحدة المعاينة هي نفسها الوحدة التجريبية، مثلاً قياس محصول القمح من قطعة أرض عولجت بسماد معين ، أو تكون وحدة المعاينة عبارة عن عينة سحبت من الوحدة التجريبية كأن نأخذ مجموعة من سنابل القمح عولجت بمبيد معين.

الخطأ التجريبي Experimental Error : هو التباين بين الوحدات التجريبية التي طبقت عليها نفس المعالجة. يتكون الخطأ التجريبي من مجموعة من العوامل غير المتحكم بها والكامنة داخل المواد التجريبية، ولا يعني هذا الخطأ وجود خطأ في التجربة وانما وجود اختلاف في المادة التجريبية نفسها، مثل الاختلاف في طبيعة كل جزء من الأرض أو الاختلاف بين الحيوانات التي أُجْرِيَتْ عليها التجربة أو في طريقة الزراعة أو الحصاد وهكذا.

ويمكن تلخيص مصادر هذا الخطأ في نقطتين هما:

1- عدم تجانس الوحدات التجريبية.

2- طريقة تنفيذ التجربة.

هذا ومن الممكن تصغير هذا الخطأ وبالتالي إعطاء التجربة مزيداً من الدقة وذلك بزيادة عدد التكرارات أو التحكم في الوحدات التجريبية وهذا ليس موضوعنا في الكتاب.

وسنتحدث هنا عن نوعين من تحليل التباين.

أولاً/ تحليل التباين الاحادي One-way ANOVA

في هذا النوع من التحليل نختبر تأثير عدة معالجات أو تأثير عامل واحد بعدة مستويات على وحدة تجريبية، وعند جمع الاستجابات من التجربة من الافضل تبويبها في جدول لسهولة فهمها والتعامل معها بإجراء عمليات حسابية عليها، فبفرض أن لدينا k معالجة وسُحِبَتْ من كل واحدة عينة بحجم

ورمزنا للاستجابة رقم j الموجودة في المعالجة i بالرمز y_{ij} فيمكن $n_i: \ i = 1,2,3, \dots, k$ كتابة الجدول كالتالي:

المعالجات	1	j	n_i	$y_{i.}$
1	y_{11} ..	y_{1j} ..	y_{1n_i}	$y_{1.}$
.	.	.	.	.
.	.	.	.	.
i	y_{i1} ..	y_{ij} ..	y_{in_i}	$y_{i.}$
.	.	.	.	.
.	.	.	.	.
k	y_{k1} ..	y_{kj} ..	y_{kn_i}	$y_{k.}$

وسنستخدم الرموز التالية لتدل على ما أُشير إليه خلال هذا الفصل:

$\boldsymbol{y_{i.}}$: مجموع الاستجابات في المعالجة i

$\boldsymbol{y_{..}}$: مجموع الاستجابات الكلي

$\boldsymbol{N}$: مجموع الاستجابات الكلي أي $N = \sum_{i=1}^{k} n_i$

$\boldsymbol{\bar{y}_{i.}}$: الوسط الحسابي للاستجابات في المعالجة i

$\boldsymbol{\bar{y}_{..}}$: الوسط الحسابي لجميع الاستجابات

نعلم من دراستنا لمساق مبادئ الاحصاء أن التباين يتناسب مع المربعات لانحرافات القيم عن وسطها الحسابي، وفي حالتنا هذه نقول أن التباين الكلي يتناسب مع مجموع المربعات التالي:

$$SSTo = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2$$

حيث يتم تحليل هذا المجموع إلى مجموع مربعات يتناسب مع التباين الذي يعزى للمعالجات ويرمز له بالرمز SST أو بالرمز $\boldsymbol{SSB}$ وهو الذي سنستخدمه في هذه الحالة من تحليل التباين و مجموع مربعات يتناسب مع التباين الذي يعزى إلى الاختلاف داخل كل معالجة ويرمز له بالرمز SSE أو بالرمز $\boldsymbol{SSW}$ وهو الذي سنستخدمه في هذه الحالة من تحليل التباين ويسمى هذا المجموع بمجموع مربعات الخطأ.

سنذكر فيما يلي عملية التحليل لمجموع المربعات الكلي:

$$SSTo = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.} + \bar{y}_{i.} - \bar{y}_{..})^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n_i}[(y_{ij} - \bar{y}_{i.}) + (\bar{y}_{i.} - \bar{y}_{..})]^2$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\bar{y}_{i.} - \bar{y}_{..})^2 + 2\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})(\bar{y}_{i.} - \bar{y}_{..})$$

لاحظ أن

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})(\bar{y}_{i.} - \bar{y}_{..}) = \sum_{i=1}^{k}(\bar{y}_{i.} - \bar{y}_{..})\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})$$

$$= (n_i\bar{y}_{i.} - n_i\bar{y}_{i.})\sum_{i=1}^{k}(\bar{y}_{i.} - \bar{y}_{..}) = 0$$

وبالتالي نحصل على

$$SSTo = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\bar{y}_{i.} - \bar{y}_{..})^2$$

حيث يعبر المجموع $\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2$ عن مجموع المربعات الذي يتناسب مع التباين الذي يعزى إلى الاختلاف داخل كل معالجة أي:

ويعبر المجموع $\sum_{i=1}^{k}\sum_{j=1}^{n_i}(\bar{y}_{i.} - \bar{y}_{..})^2$ عن مجموع المربعات الذي يتناسب مع التباين الذي يعزى للمعالجات.

وبإجراء عمليات حسابية بسيطة نحصل على الصيغ التالية و وهي التي سنستخدمها في حساباتنا:

$$\boldsymbol{SSTo = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij})^2 - CF}$$

$$\boldsymbol{SSB = \sum_{i=1}^{k}\frac{(y_{i.})^2}{n_i} - CF}$$

حيث $\boldsymbol{CF = \frac{(y_{..})^2}{N}}$ والتي تسمى معامل التصحيح **Correction Factor**

ويكون:

$$\boldsymbol{SSW = SSTo - SSB}$$

وتكون درجات الحرية كالتالي:

SS	df
SSB	$k-1$
SSW	$N-k$
$SSTo$	$N-1$

بعد ايجاد المجاميع السابقة نكون ما يسمى بجدول التباين كما يلي:

جدول التباين الاحادي

مصدر الاختلاف S.O.V	درجة الحرية df	مجموع المربعات SS	متوسط المربعات(التباين) MS	قيمة F_{cal} المحسوبة
بين المعالجات	$k-1$	SSB	$MSB=\frac{SSB}{k-1}$	$F_{cal}=\frac{MSB}{MSW}$
الخطأ التجريبي	$N-k$	SSW	$MSW=\frac{SSW}{N-k}$	
المجموع	$N-1$	$SSTO$		

بعد تكوين جدول التباين نكون قد أوجدنا من قبل قيمة $F_{(\alpha,k-1,N-k)}$ الجدولية حسب مستوى المعنوية المطلوب ثم نقارن بين قيمة F_{cal} المحسوبة وقيمتها الجدولية مع الاخذ في الاعتبار أن الاختبار ذو اتجاه واحد من اليمين وأن الفرضيات هي:

$$H_0: \mu_1=\mu_2=\mu_3=\cdots\cdots=\mu_k$$

H_1:على الاقل أحد الأوساط مختلف

مثال5.1: يمثل الجدول التالي نتائج طلبة تم تدريسهم مادة الاحصاء التطبيقي بأربعة أساليب مختلفة. اختبر وجود فرق بين الأساليب الأربعة عند مستوى معنوية 0.05

الأسلوب الأول T_1	الأسلوب الثاني T_2	الأسلوب الثالث T_3	الأسلوب الرابع T_4
65	75	59	94
87	69	78	89
73	83	67	80
79	81	62	88
81	72	83	
69	79	76	
	90		

الحل/

هنا يمثل كل أسلوب معالجة، فنكون الجدول الخاص بإيجاد المجاميع أولاً ثم نوجد مجاميع المربعات ثم نكون جدول التباين.

المعالجة	1	2	3	4	5	6	7	$y_{i.}$
T_1	65	87	73	79	81	69		454
T_2	75	69	83	81	72	79	90	549
T_3	59	78	67	62	83	76		425
T_4	94	89	80	88				351
								$y_{..} = 1779$

$$N = 6 + 7 + 6 + 4 = 23$$

$$CF = \frac{(y_{..})^2}{N} = \frac{(1779)^2}{23} = 137601.783$$

$$SSTo = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij}\right)^2 - CF = (65)^2 + (87)^2 + \cdots + (88)^2 - 137601.783$$

$$= 139511 - 137601.783 = 1909.217$$

$$SSB = \sum_{i=1}^{k}\frac{(y_{i.})^2}{n_i} - CF$$

$$= \frac{(454)^2}{6} + \frac{(549)^2}{7} + \frac{(425)^2}{6} + \frac{(351)^2}{4} - 137601.783 = 712.586$$

$SSW = SSTo - SSB = 1909.217 - 712.586 = 1196.631$

الفرضيات:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$$

H_1: على الأقل أحد الأوساط مختلف

القيم الحرجة المعيارية:

حيث أن مستوى المعنوية 0.05 فإن $F_{(0.05,3,19)} = 3.1274$

المعيار:

جدول التباين الأحادي

S.O.V	df	SS	MS	F_{cal}
Between Treatments	3	712.586	237.529	$F_{cal} = 3.7714$
Within Treatments (Error)	19	1196.631	62.981	
Total	22	1909.217		

المقارنة و القرار:

وحيث أن $(F_{cal} = 3.7714) > (F_{(0.05,3,19)} = 3.1274)$ فإننا نرفض الفرضية الصفرية وهذا يعني أن أوساط أساليب التدريس ليست متساوية وبالتالي نستنتج أن هناك فرقاً بين أساليب التدريس الاربعة .

من المنطق أن نتساءل بعد حل المثال السابق ورفض الفرضية الصفرية عن الأسلوب أو الأساليب التي يختلف وسطها الحسابي وذلك على الأقل لاستبعاده واستخدام الأساليب التي أُثْبِتَتْ فعاليتها، وهذ يقودنا لدراسة اختبارات جديدة تسمى اختبارات المقارنات المتعددة.

اختبارات المقارنات المتعددة Multiple Comparison Tests

هناك اختبارات كثيرة تستخدم للمقارنات المتعددة مثل:

1- **اختبار أقل فرق معنوي Least Significance Difference Test (L.S.D.)**

2- **اختبار دونيت Dunnett's Test**

3- **اختبارات دنكان – توكي – شيفيه وغيرها**

وسنتحدث هنا عن أول اختبارين فقط.

1- اختبار أقل فرق معنوي (L.S.D.)

يُعَدُ هذا الاختبار من أفضل الاختبارات وذلك لدقته وسهولته، ولكي نستخدم هذا الاختبار لابد أن تكون الفرضية الصفرية قد رُفِضَتْ، أي هناك فروق بين المتوسطات.

معيار الاختبار:

$$\mathbf{L.S.D} = \boldsymbol{t}_{\left(\frac{\alpha}{2},\vartheta\right)} \times \sqrt{\boldsymbol{MSE} \times \left(\frac{\mathbf{1}}{\boldsymbol{n_i}} + \frac{\mathbf{1}}{\boldsymbol{n_j}}\right)}$$

حيث α هي المعنوية المحسوبة في تحليل التباين، MSE متوسط مربعات الخطأ التجريبي و ϑ هي درجة الحرية للخطأ التجريبي.

ولكي نتحقق من معنوية الفرق **نوجد جميع الفروقات المطلقة بين متوسطات المعالجات ونقارنها بقيمة الاختبار L.S.D** وكل فرق بين معالجتين يكون أكبر من قيمة الاختبار دلَّ ذلك على وجود فرق معنوي بين هاتين المعالجتين.

مثال5.2: في مثال 5.1 استخدم اختبار اقل فرق معنوي **L.S.D** لتحديد نوع الفرق بين متوسطات اساليب التدريس التي استخدمت.

الحل/

نحسب المعيار للاختبار كما يلي:

$$\mathrm{L.S.D} = t_{\left(\frac{\alpha}{2},\vartheta\right)} \times \sqrt{MSE \times \left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

$$= t_{(0.025,19)} \times \sqrt{62.981 \times \left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

$$= 2.093 \times \sqrt{62.981 \times \left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

يتم حساب قيمة $\mathrm{L.S.D}$ لكل فرق كما يلي:

$$\mathrm{L.S.D}(\mathrm{T}_1,\mathrm{T}_2) = 2.093 \times \sqrt{62.981 \times \left(\frac{1}{6}+\frac{1}{7}\right)} = 9.241$$

$$\mathrm{L.S.D}(\mathrm{T}_1,\mathrm{T}_3) = 2.093 \times \sqrt{62.981 \times \left(\frac{1}{6}+\frac{1}{6}\right)} = 9.59$$

وهكذا لبقية الفروق ثم نكون الجدول التالي:

المعالجة	n_i	$y_{i.}$	$\bar{y}_{i.}$	$\lvert\bar{y}_{i.}-\bar{y}_{j.}\rvert$		$\mathrm{L.S.D}$	يوجد فرق معنوي
T_1	6	454	75.667	$\lvert\bar{y}_{1.}-\bar{y}_{2.}\rvert$	2.762	9.241	لا
T_2	7	549	78.429	$\lvert\bar{y}_{1.}-\bar{y}_{3.}\rvert$	4.834	9.59	لا
T_3	6	425	70.833	$\lvert\bar{y}_{1.}-\bar{y}_{4.}\rvert$	12.083	10.722	نعم
T_4	4	351	87.75	$\lvert\bar{y}_{2.}-\bar{y}_{3.}\rvert$	7.596	9.241	لا
				$\lvert\bar{y}_{2.}-\bar{y}_{4.}\rvert$	9.321	10.411	لا
				$\lvert\bar{y}_{3.}-\bar{y}_{4.}\rvert$	16.917	10.722	نعم

نلاحظ من الجدول السابق وبمقارنة الفروق المطلقة بين متوسطات المعالجات وبين قيمة $\mathrm{L.S.D}$ المقابلة لها نلاحظ وجود فرق معنوي بين أسلوبي التدريس الأول و الرابع وبين أسلوبي التدريس الثالث والرابع.

2- اختبار دونيت **Dunnett's Test**

يتميز اختبار دونيت بإمكانية تطبيقه بغض النظر عن معنوية F وذلك لأن المقارنات التي يختبرها تكون محددة سلفاً قبل اجراء التجربة، ويشترط اختبار دونيت تساوي حجوم العينات في التجربة.

ويعتبر اختبار دونيت تعديلاً لاختبار t الخاص بمقارنة وسطين لمجتمعين طبيعيين مستقلين، ويتم تحديد معالجة معينة واعتبارها معالجة سيطرة Control Treatment حيث يتم مقارنة بقية المعالجات بها، لذلك تأخذ الفرضيات الشكل التالي:

$$H_0: \mu_i = \mu_0 \qquad , i = 1,2,\cdots\cdots,k-1$$
$$H_1: \mu_i \neq \mu_0 \qquad , i = 1,2,\cdots\cdots,k-1$$

حيث μ_0 متوسط معالجة السيطرة.

معيار الاختبار:

$$\mathbf{D} = \boldsymbol{d}_{(\alpha,\vartheta,k)} \times \sqrt{\frac{\mathbf{2} \times \boldsymbol{MSE}}{\boldsymbol{n}}}$$

حيث α هي المعنوية المحسوبة في تحليل التباين، ϑ هي درجة الحرية للخطأ التجريبي، k عدد المعالجات، $\boldsymbol{n}$ حجم العينة، هذا ويتم ايجاد قيمة $\boldsymbol{d}_{(\alpha,\vartheta,k)}$ باستخدام جدول دونيت.

حيث يتم رفض H_0 عندما يكون $|\bar{\boldsymbol{y}}_{i.} - \bar{\boldsymbol{y}}_{\mathbf{0}}| > \mathbf{D}$

مثال5.3: الجدول التالي يُبَيّن الاستجابات لعينات حجمها 3 أخذت لأربعة معالجات، حدد المعالجات التي تختلف معنوياً عند مستوى معنوية 0.05 على اعتبار أن المعالجة الثالثة هي معالجة السيطرة.

المعالجات	y_{ij}			$y_{i.}$
T_1	5	5	4	14
T_2	4	3	3	10
T_3	0	1	0	1
T_4	1	2	2	5

الحل/في هذه الحالة تكون الفرضيات:

$$H_0: \mu_i = \mu_3 \qquad , i = 1,2,4$$
$$H_1: \mu_i \neq \mu_3 \qquad , i = 1,2,4$$

وحيث أن حساب معيار الاختبار يتطلب معرفة $\boldsymbol{MSE}$ سنقوم بإنشاء جدول التباين الخاص بالتجربة، وبعد ايجاد مجاميع المربعات كما تعلمنا نجد أن:

جدول التباين الأحادي

S.O.V	df	SS	MS	F_{cal}
Between Treatments	3	32.33	10.787	32.27
Within Treatments (Error)	8	2.67	0.334	
Total	11	35		

من جدول التباين نجد أن $MSE = 0.334$ ومن جدول دونيت نجد أن $d_{(0.05,8,4)} = 2.88$

$$D = 2.88 \times \sqrt{\frac{2 \times 0.334}{3}} = 1.36$$

ويكون الوسط الحسابي لمعالجة السيطرة $\bar{y}_0 = \bar{y}_{3.} = 0.33$

المعالجة	n_i	$y_{i.}$	$\bar{y}_{i.}$	$\lvert\bar{y}_{i.} - \bar{y}_{j.}\rvert$		D	يوجد فرق معنوي
T_1	3	14	4.67	$\lvert\bar{y}_{1.} - \bar{y}_0\rvert$	4.34	1.36	نعم
T_2	3	10	3.33	$\lvert\bar{y}_{2.} - \bar{y}_0\rvert$	3	1.36	نعم
T_4	3	5	1.67	$\lvert\bar{y}_{4.} - \bar{y}_0\rvert$	0.92	1.36	لا

نلاحظ من الجدول السابق وبمقارنة الفروق المطلقة بين المتوسطات وبين المعيار D نجد وجود فروق بين متوسطات كل من المعالجة الأولى و الثانية ومعالجة السيطرة.

مثال5.4: الجدول التالي يبين نتائج اجراء تجربة على 5 معالجات مستخدمين عينة من 8 مشاهدات والمطلوب استخدام اختبار دونيت عند مستوى معنوية 0.05 لتحديد المعالجات المختلفة معنوياً باعتبار أن T_3 هي معالجة السيطرة.

	T_1	T_2	T_3	T_4	T_5
	3	2	14	29	24
	5	12	6	20	26
	1	13	12	36	40
	8	6	4	21	32
	1	10	19	25	20
	1	7	3	18	33
	4	11	9	26	27
	9	19	21	17	30
المعدل	4	10	11	24	29

الحل/

في هذه الحالة تكون الفرضيات:

$$H_0: \mu_i = \mu_3 \qquad , i = 1,2,4,5$$
$$H_1: \mu_i \neq \mu_3 \qquad , i = 1,2,4,5$$

وحيث أن حساب معيار الاختبار يتطلب معرفة MSE سنقوم بإنشاء جدول التباين الخاص بالتجربة، وبعد ايجاد مجاميع المربعات كما تعلمنا نجد أن:

جدول التباين الاحادي

S.O.V	df	SS	MS	F_{cal}
Treatments	4	3497.3	874.4	27.33
Error	35	1120	32	
Total	39	4617.6		

من جدول التباين نجد أن $MSE = 32$ ومن جدول دونيت نجد أن $d_{(0.05,35,5)} = 2.56$

$$D = 2.56 \times \sqrt{\frac{2 \times 32}{8}} = 7.24$$

ويكون الوسط الحسابي لمعالجة السيطرة $\bar{y}_0 = \bar{y}_{3.} = 11$

المعالجة	n_i	$\bar{y}_{i.}$	$\lvert \bar{y}_{i.} - \bar{y}_{j.} \rvert$		D	يوجد فرق معنوي
T_1	8	4	$\lvert \bar{y}_{1.} - \bar{y}_0 \rvert$	7	7.24	لا
T_2	8	10	$\lvert \bar{y}_{2.} - \bar{y}_0 \rvert$	1	7.24	لا
T_4	8	24	$\lvert \bar{y}_{4.} - \bar{y}_0 \rvert$	13	7.24	نعم
T_5	8	29	$\lvert \bar{y}_{5.} - \bar{y}_0 \rvert$	18	7.24	نعم

ثانياً/ تحليل التباين الثنائي Two-way ANOVA

في هذا النوع من التحليل نختبر تأثير عاملين كلٍ بعدة مستويات وقد يكون العاملان مستقلان لا يحدث بينهما تفاعل في التجربة أو غير مستقلان فيحدث تفاعل بينهما في التجربة.

وسندرس الحالتين التاليتين:

- **تحليل التباين الثنائي بدون تداخل Two-way ANOVA (without interaction)**
- **تحليل التباين الثنائي بالتداخل Two-way ANOVA (with interaction)**

أولاً: تحليل التباين الثنائي (بدون تداخل)

بفرض أن لدينا عاملين A و B مستقلين عن بعضهما البعض بحيث كان العامل الأول بعدد a من المستويات وهي $\{A_1, A_2, \dots, A_a\}$ والعامل الثاني بعدد b من المستويات وهي $\{B_1, B_2, \dots, B_b\}$ وقمنا بجمع الاستجابات من التجربة، ومن ثم سيتم تبويبها في جدول لسهولة فهمها والتعامل معها بإجراء عمليات حسابية عليها كالتالي:

A \ B	B_1	$\cdots$	B_j	$\cdots$	B_b	المجموع
A_1	y_{11}	$\vdots$	y_{1j}	$\vdots$	y_{1b}	$y_{1.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_i	y_{i1}	$\vdots$	y_{ij}	$\vdots$	y_{ib}	$y_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_a	y_{aj}	$\vdots$	y_{aj}	$\vdots$	y_{ab}	$y_{a.}$
المجموع	$y_{.1}$	$\cdots$	$y_{.j}$	$\cdots$	$y_{.b}$	$y_{..}$

حيث تمثل الصفوف مستويات العامل A والاعمدة مستويات العامل B.

وتكون الفرضيات التي يمكن اختبارها هي:

$$H_0^1: \mu_{A1} = \mu_{A2} = \mu_{A3} = \cdots\cdots = \mu_{Aa}$$
$$H_0^2: \mu_{B1} = \mu_{B2} = \mu_{B3} = \cdots\cdots = \mu_{Bb}$$

مقابل الفرضيات البديلة:

H_1^1: A على الأقل أحد الأوساط مختلف للعامل

H_1^2: B على الأقل أحد الأوساط مختلف للعامل

وتكون مجاميع المربعات كما يلي:

$$SSTo = \sum_{i=1}^{b}\sum_{j=1}^{a}(y_{ij})^2 - CF$$

$$SSA = \frac{\sum_{i=1}^{a}(y_{i.})^2}{b} - CF$$

$$SSB = \frac{\sum_{j=1}^{b}(y_{.j})^2}{a} - CF$$

$$SSE = SSTo - (SSA + SSB)$$

حيث معامل التصحيح $CF = \frac{(y_{..})^2}{ab}$

ويكون جدول التباين في هذه الحالة:

جدول التباين الثنائي بدون تداخل

S.O.V	df	SS	MS	F_{cal}
بين الصفوف مستويات A	$a-1$	SSA	$MSA = \frac{SSA}{a-1}$	$F_A = \frac{MSA}{MSE}$
بين الاعمدة مستويات B	$b-1$	SSB	$MSB = \frac{SSB}{b-1}$	$F_B = \frac{MSB}{MSE}$
الخطأ التجريبي	$(a-1)(b-1)$	SSE	$MSE = \frac{SSE}{(a-1)(b-1)}$	
المجموع	$ab-1$	$SSTO$		

وتكون القيم الحرجة المعيارية في هذه الحالة كما يلي:

$F_{(\alpha,a-1,(a-1)(b-1))}$ نقارن بينها وبين قيمة F_A لاختبار تأثير العامل A.

$F_{(\alpha,b-1,(a-1)(b-1))}$ نقارن بينها وبين قيمة F_B لاختبار تأثير العامل B.

مع الأخذ في الاعتبار أن الاختبار ذو اتجاه واحد من اليمين .

مثال5.5: أراد مهندس زراعي اختبار معنوية الفرق بين ثلاثة مستويات تركيز لنوع جديد من السماد واختبار معنوية الفرق بين 4 أنواع من البذور عند مستوى معنوية 0.05 ، فلجأ إلى إعداد 12 قطعة أرض متجانسة وزرعت كل قطعة منها بمستوى تركيز مع نوع من البذور، وبعد انتهاء التجربة كانت نتائج الحصاد بالكيلوجرام كما يلي:

انواع البذور B / تركيز السماد A	B_1	B_2	B_3	B_4	$y_{i.}$
A_1	12	9	8	10	39
A_2	13	13	10	11	47
A_3	15	14	10	10	49
$y_{.j}$	40	36	28	31	135

الحل/

لاحظ أن $b = 4 , a = 3$

معامل التصحيح $CF = \frac{(y_{..})^2}{ab} = \frac{(135)^2}{3\times4} = 1518.75$

$$SSTo = \sum_{i=1}^{3}\sum_{j=1}^{4}(y_{ij})^2 - CF = [12^2 + 13^2 + \cdots + 10^2] - 1518.75 = 50.25$$

$$SSA = \frac{\sum_{i=1}^{3}(y_{i.})^2}{b} - CF = \frac{39^2 + 47^2 + 49^2}{4} - 1518.75 = 14$$

$$SSB = \frac{\sum_{j=1}^{4}(y_{.j})^2}{a} - CF = \frac{40^2 + 36^2 + 28^2 + 31^2}{3} - 1518.75 = 28.25$$

$$SSE = SSTo - (SSA + SSB) = 50.25 - (14 + 28.25) = 8$$

الفرضيات:

$$H_0^1: \mu_{A1} = \mu_{A2} = \mu_{A3}$$
$$H_0^2: \mu_{B1} = \mu_{B2} = \mu_{B3} = \mu_{B4}$$

H_1^1: A على الأقل أحد الأوساط مختلف للعامل

H_1^2: B على الأقل أحد الأوساط مختلف للعامل

القيم الحرجة المعيارية:

حيث أن مستوى المعنوية 0.05 فإن:

القيمة الجدولية لاختبار العامل A هي $F_{(0.05,2,6)} = 5.14$

القيمة الجدولية لاختبار العامل B هي $F_{(0.05,3,6)} = 4.76$

المعيار:

جدول التباين

S.O.V	df	SS	MS	F_{cal}
A	2	14	7	$\mathbf{F_A = 5.25}$
B	3	28.25	9.417	$\mathbf{F_B = 7.06}$
E	6	8	1.333	
المجموع	11	1518.75		

المقارنة و القرار:

وحيث أن $(F_{(0.05,2,6)} = 5.14) < (F_A = 5.25)$ فإننا نرفض الفرضية الصفرية $H_0(A)$ وهذا يعني أن مستويات تركيز السماد مختلفة معنوياً عن بعضها البعض.

وحيث أن $(F_{(0.05,3,6)} = 4.76) < (F_B = 7.06)$ فإننا نرفض الفرضية الصفرية $H_0(B)$ وهذا يعني أن أنواع البذور مختلفة معنوياً عن بعضها البعض.

ثانياً: تحليل التباين الثنائي (بالتداخل)

بشكل عام نجد أنه إذا اشترك عاملين في تجربة واحدة فإنه ينتج تفاعل بين العاملين وقد نكون مهتمين بهذا التفاعل حيث أنه يعني وجود مصدر آخر للاختلاف، ويكون النموذج الرياضي الذي يوضح تأثير العوامل وتداخلها في تحديد قيمة الاستجابة:

$$\boldsymbol{y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}}$$

حيث:

$\boldsymbol{\mu}$: قيمة متوسط المجتمع ويقدر بمتوسط قيم التجربة.

$\boldsymbol{\alpha_i}$: تأثير المستوى $\boldsymbol{i}$ من العامل الأول حيث سنفرض أن $\sum \alpha_i = 0$.

$\boldsymbol{\beta_j}$: تأثير المستوى $\boldsymbol{j}$ من العامل الثاني حيث سنفرض أن $\sum \beta_i = 0$.

$\boldsymbol{\gamma_{ij}}$: تأثير تفاعل المستوى i من العامل الأول و المستوى j من العامل الثاني حيث سنفرض أن $\sum \sum \boldsymbol{\gamma_{ij}} = 0$.

$\boldsymbol{\varepsilon_{ijk}}$: قيمة الخطأ التجريبي وتكون مستقلة عن بعضها البعض ، حيث سنفرض أن:

$$\sum \sum \sum \boldsymbol{\varepsilon_{ijk}} = 0.$$

وسنفرض أيضاً أن:

$$\boldsymbol{\alpha_i , \beta_j , \varepsilon_{ijk} \sim N(0, \sigma^2)}$$

ويكون جدول الاستجابات التي نحصل عليها من التجربة كالتالي:

A \ B	B_1	$\cdots$	B_j	$\cdots$	B_b	المجموع
A_1	y_{111} y_{112} $\vdots$ y_{11n}	$\vdots$	y_{1j1} y_{1j2} $\vdots$ y_{1jn}	$\vdots$	y_{1b1} y_{1b2} $\vdots$ y_{1bn}	$y_{1..}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_i	y_{i11} y_{i12} $\vdots$ y_{i1k} $\vdots$ y_{i1n}	$\vdots$	y_{ij1} y_{ij2} $\vdots$ y_{ijk} $\vdots$ y_{ijn}	$\vdots$	y_{ib1} y_{ib2} $\vdots$ y_{ibk} $\vdots$ y_{ibn}	$y_{i..}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_a	y_{a11} y_{a12} $\vdots$ y_{a1n}	$\vdots$	y_{1j1} y_{1j2} $\vdots$ y_{ajn}	$\vdots$	y_{1b1} y_{1b2} $\vdots$ y_{abn}	$y_{a..}$
المجموع	$y_{.1.}$	$\cdots$	$y_{.j.}$	$\cdots$	$y_{.b.}$	$y_{...}$

حيث تمثل الصفوف مستويات العامل A وعددها a والأعمدة مستويات العامل B وعددها b وتمثل n عدد الاستجابات التي عُولِجَت بالمستوى i من العامل الأول و المستوى j من العامل الثاني .

وتكون الفرضيات التي يمكن اختبارها هي:

$$H_0^1: \mu_{A1} = \mu_{A2} = \mu_{A3} = \cdots\cdots = \mu_{Aa}$$
$$H_0^2: \mu_{B1} = \mu_{B2} = \mu_{B3} = \cdots\cdots = \mu_{Bb}$$
$$H_0^3: \mu_{AB1} = \mu_{AB2} = \mu_{AB3} = \cdots\cdots = \mu_{ABb}$$

أو

$$H_0^1: \alpha_1 = \alpha_2 = \alpha_3 = \cdots\cdots = \alpha_a$$
$$H_0^2: \beta_1 = \beta_2 = \beta_3 = \cdots\cdots = \beta_a$$
$$H_0^3: \gamma_{11} = \gamma_{12} = \gamma_{13} = \cdots\cdots = \gamma_{ab}$$

مقابل الفرضيات البديلة:

على الأقل أحد الأوساط مختلف للعامل A :H_1^1

على الأقل أحد الأوساط مختلف للعامل B :H_1^2

على الأقل أحد الأوساط مختلف للتداخل AB :H_1^3

وتكون مجاميع المربعات كما يلي:

$$SSTo = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}\right)^2 - CF$$

$$SSA = \frac{\sum_{i=1}^{a}(y_{i..})^2}{bn} - CF$$

$$SSB = \frac{\sum_{j=1}^{b}\left(y_{.j.}\right)^2}{an} - CF$$

$$SSAB = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\left(y_{ij.}\right)^2}{n} - SSA - SSB - CF$$

$$SSE = SSTo - (SSA + SSB + SSAB)$$

حيث معامل التصحيح $CF = \frac{(y_{...})^2}{abn}$

ويكون جدول التباين في هذه الحالة:

جدول التباين الثنائي بدون تداخل

S.O.V	df	SS	MS	F_{cal}
بين مستويات العامل A	$a-1$	SSA	$MSA = \frac{SSA}{a-1}$	$F_A = \frac{MSA}{MSE}$
بين مستويات العامل B	$b-1$	SSB	$MSB = \frac{SSB}{b-1}$	$F_B = \frac{MSB}{MSE}$
التفاعل AB	$(a-1)(b-1)$	$SSAB$	$MSAB = \frac{SSAB}{(a-1)(b-1)}$	$F_{AB} = \frac{MSAB}{MSE}$
الخطأ التجريبي	$ab(n-1)$	SSE	$MSE = \frac{SSE}{ab(n-1)}$	
المجموع	$abn-1$	$SSTO$		

وتكون القيم الحرجة المعيارية في هذه الحالة كما يلي:

$F_{(\alpha,a-1,ab(n-1))}$ نقارن بينها وبين قيمة F_A لاختبار تأثير العامل A.

$F_{(\alpha,b-1,ab(n-1))}$ نقارن بينها وبين قيمة F_B لاختبار تأثير العامل B.

$F_{(\alpha,(a-1)(b-1),ab(n-1))}$ نقارن بينها وبين قيمة F_{AB} لاختبار التفاعل بين العاملين.

مع الاخذ في الاعتبار أن الاختبار ذو اتجاه واحد من اليمين .

مثال5.6: اختبر معنوية الفرق بين أسلوبي التدريب ومعنوية الفرق بين الفئات العمرية للشباب، وكذلك اختبر تداخل هذين العاملين عند مستوى معنوية 0.05 إذا علمت أن درجات تقييم المدرب المشرف على الشباب كانت كما هو موضح بالجدول التالي:

فئات العمر B / أسلوب التدريب A	B_1	B_2	B_3	B_4	$y_{i..}$
A_1	6 6 7	7 7 6	5 6 7	8 9 8	82
A_2	7 7 6	6 6 8	8 9 10	7 8 7	89
$y_{.j.}$	39	40	45	47	171

الحل/

لاحظ أن $n = 3\ ,\ \ b = 4\ , a = 2$

معامل التصحيح $CF = \frac{(y_{...})^2}{ab} = \frac{(171)^2}{2\times4\times3} = 1218.375$

$$SSTo = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}\right)^2 - CF$$

$$= [6^2 + 6^2 + \cdots + 7^2] - 1218.375 = 32.625$$

$$SSA = \frac{\sum_{i=1}^{a}(y_{i..})^2}{bn} - CF = \frac{82^2 + 89^2}{4\times3} - 1218.375 = 2.042$$

$$SSB = \frac{\sum_{j=1}^{b}(y_{.j.})^2}{an} - CF = \frac{39^2 + 40^2 + 45^2 + 47^2}{2 \times 3} - 1218.375 = 7.458$$

لإيجاد قيم $y_{ij.}$ نكون الجدول التالي:

فئات العمر B / اسلوب التدريب A	B_1	B_2	B_3	B_4
A_1	19	20	18	25
A_2	20	20	27	22

$$SSAB = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij.})^2}{n} - SSA - SSB - CF$$
$$= \frac{19^2 + 20^2 + \cdots + 22^2}{3} - 2.042 - 7.458 - 1218.375 = 13.125$$

$$SSE = SSTo - (SSA + SSB + SSAB)$$
$$= 32.625 - (2.042 + 7.458 + 13.125) = 10$$

الفرضيات:

$$H_0^1: \alpha_1 = \alpha_2$$
$$H_0^2: \beta_1 = \beta_2 = \beta_3 = \beta_4$$
$$H_0^3: \gamma_{11} = \gamma_{12} = \gamma_{13} = \cdots\cdots = \gamma_{24}$$

مقابل الفرضيات البديلة:

H_1^1: على الأقل أحد الأوساط مختلف للعامل A

H_1^2: على الأقل أحد الأوساط مختلف للعامل B

H_1^3: على الأقل أحد الأوساط مختلف للتداخل AB

القيم الحرجة المعيارية:

حيث أن مستوى المعنوية 0.05 فإن:

القيمة الجدولية لاختبار العامل A هي $F_{(0.05,1,16)} = 4.49$

القيمة الجدولية لاختبار العامل B هي $F_{(0.05,3,16)} = 3.24$

القيمة الجدولية لاختبار التفاعل بين العاملين هي $F_{(0.05,3,16)} = 3.24$

المعيار:

جدول التباين

S.O.V	df	SS	MS	F_{cal}
A	1	2.042	2.042	$F_A = 3.27$
B	3	7.458	2.486	$F_B = 4$
AB	3	13.125	4.375	$F_{AB} = 7$
Error	16	10	0.625	
المجموع	23	32.625		

المقارنة و القرار:

وحيث أن $(F_{(0.05,1,16)} = 4.49) < (F_A = 3.27)$ فإننا نقتل الفرضية الصفرية $H_0(A)$ وهذا يعني أنه لا فرق بين اسلوبي التدريب.

وحيث أن $(F_{(0.05,3,16)} = 3.24) > (F_B = 4)$ فإننا نرفض الفرضية الصفرية $H_0(B)$ وهذا يعني أنه يوجد على الاقل فئة عمرية تختلف معنوياً.

وحيث أن $(F_{(0.05,3,16)} = 3.24) > (F_{AB} = 7)$ فإننا نرفض الفرضية الصفرية $H_0(AB)$ وهذا وجود اختلاف معنوي بين مستويات تداخل العاملين.

تمارين:

1- في مركز تدريب، تم تقسيم طلبة دورة تخصصية إلى ثلاثة مجموعات متجانسة. وقد تلقى الطلبة المحاضرات عن طريق ثلاث محاضرين من بلاد مختلفة بحيث يكون لكل مجموعة محاضر معين، وبعد الانتهاء من الدورة كانت النتائج كما يلي:

المحاضر	درجة الطالب					
I	5	7	6	8	8	7
II	6	8	7	7	6	6
III	6	9	8	7	9	6

المطلوب عند مستوى معنوية 0.05 :

- اختبار معنوية الفروق بين المجموعات.
- اجراء اختبار أقل فرق معنوي.
- اجراء اختبار دونيت باعتبار أن المعالجة III هي معالجة السيطرة.

2- قسمت أرض زراعية إلى 15 قطعة متجانسة ومن ثم تم زراعتها بأربعة أصناف من الفاصوليا فأعطت النتائج التالية:

صنف الفاصوليا	كمية المحصول بالكيلوجرام				
A	80	75	80	70	
B	90	90	80		
C	65	65	70	80	75
D	90	95	85		

المطلوب عند مستوى معنوية 0.05 :

- اختبار معنوية الاختلاف بين أصناف الفاصوليا.
- تحديد نوعية الفروق بين الصنف A وكل من الصنفين C,B

3- أكمل جدول تحليل التباين التالي ثم اختبر فرضية العدم عند معنوية 0.05 :

$$H_0: \mu_1 = \mu_2 = \mu_3 = \cdots\cdots = \mu_5$$

مصدر الاختلاف	درجة الحرية	مجموع المربعات	التباين	F
بين المعالجات		20		
الخطأ التجريبي				
الاختلاف الكلي	20	50		

4- اختبر معنوية الفروق على دلالة 0.05 بين مستويات العامل الاول ومستويات العامل الثاني لنتائج التجربة التالية، حيث تمثل الاستجابات سرعة السيارة في الساعة (كم/ساعة).

نوع البنزين / نوع السيارة	III	II	I
A	131	140	125
B	133	135	130
C	129	133	127
D	134	136	128

5- اختبر معنوية الفروق بين مستويات العامل الاول ومستويات العامل الثاني وذلك بعد اكمال جدول التباين وذلك عند مستوى معنوية 0.05

S.O.V	df	SS	MS	F
A				
B		28	9.333	
Error	12		2.83	
Total	19	100		

6- تم تصنيف مجموعة من 18 طالب في المرحلة الثانوية وفقاً لأطوالهم و أوزانهم بهدف اختبار قدراتهم في إنجاز فعالية رياضية معينة، فكانت نتائج التجربة(زمن الانجاز بالثانية) كالتالي:

الطول \ الوزن بالكجم	50-55	55-60	60-65	65-70
قصير القامة	30 31	33 34	25 37	39 40
طويل القامة	27 26	26 26	30 31	30 28

اختبر عند 0.01 معنوية الفروق بين مستويات العامل الاول ومستويات العامل الثاني وتفاعلهما.

7- تم قياس كمية الشوائب في خام أحد المعادن المستخرج من ثلاث مناجم بفحص 5 عينات من كل منجم فكانت القياسات كما يلي:

المنجم		
C	B	A
7	6.5	6.6
6.9	6.7	6.2
6.8	6.5	6.4
7.1	6.7	6.3
6.9	6.6	6.4

عند مستوى معنوية 0.05 اجب عن ما يلي:

- هل هناك دليلاً كافياً على اختلاف كمية الشوائب في خامات المناجم الثلاثة.
- حدد أياً من المعادن ذو نوعية أفضل، هل هو المستخرج من المنجم A أم المستخرج من المنجم C .

8- الجدول التالي يبين الاستجابات التي حصل عليها أحد الباحثين من تجربة على عاملين لمعرفة تأثير كل عامل و التفاعل بينهما:

A \ B	B_1	B_2	B_3	المجموع
A_1	7 33 26 27 21 6 14 19	6 11 18 14 19 14	9 12 6 24 1 10	297
	153	82	62	
A_2	42 8 28 30 22 17 32	28 6 1 2 37	13 10 1 6 10	293
	179	74	40	
المجموع	332	156	102	590

اختبر عند مستوى معنوية 0.05 وجود فروق معنوية بين مستويات العامل الاول ومستويات العامل الثاني والتفاعل بين المستويين.

الفصل السادس

الاختبارات غير المعلمية

Non-parametric Tests

تسمى الاختبارات التي درسناها في الفصول السابقة بالاختبارات المعلمية **Parametric Tests** وذلك لأنها تتعلق باختبار فرضيات حول معلمة من معالم المجتمع مثل اختبارات حول وسط المجتمع أو الفرق بين متوسطي مجتمعين وكذلك اختبارات حول النسبة و التباين، وتعلمنا أنه لابد من تطبيق شروط معينة حتى نتمكن من تطبيق الاختبارات مثل ضرورة كون المجتمعات التي نختبر أوساطها طبيعية أو قريبة جداً من كونها طبيعية و من هنا نتساءل: ماذا لو طبقنا الاختبارات دون تحقق الشروط الواجب توفرها لتطبيقها؟ والاجابة هي اننا قد نحصل على نتائج غير سليمة وبالتالي نتخذ قرارات خاطئة.

ومن ناحية اخرى قد لا نستطيع الحصول على قياسات عددية دقيقة للظواهر المدروسة، ولكن من الممكن الحصول على ترتيب لقيم هذه الظواهر، كأن نحصل على تقديرات لفظية لدرجات الطلبة في امتحان معين، أو ترتيب مجموعة من الطلبة حسب انجازهم نشاط معين وهكذا..

لذا دعت الحاجة لإيجاد طرق لتحليل هذه البيانات لا تتطلب شروطاً لتطبيقها، وتسمى هذه الطرق اختبارات غير معلميه.

ان لتطبيق هذه الاختبارات مزايا عدة، فمن ناحية تعتبر سهلة الاستخدام وتعطي نتائج بسرعة ومفيدة في حالة تعذر تطبيق الاختبارات المعلمية. ولكن تطبيق هذه الاختبارات يوقع بعض المضار فمثلاً استخدامها عند توفر شروط الاختبارات المعلمية يؤدي لفقدان بعض المعلومات الموجودة في البيانات حيث أننا نستخدم رتب البيانات ونهمل قيم البيانات نفسها، وغالباً لا تجرى هذه الاختبارات عند مستوى معنوية معين كما الحال في التوزيعات المنفصلة.

مما سبق فإننا يمكن أن نستخدم الاختبارات غير المعلمية في الحالات التالية:

1- في اختبار فرضيات لا تتعلق بمعالم المجتمع.
2- لا تتوفر الشروط اللازمة لتطبيق الاختبارات المعلمية.
3- اذا كانت البيانات المتوفرة لدينا غير مناسبة للاختبارات المعلمية.
4- الحصول على قرار سريع.

وسنتحدث عن بعض الاختبارات الهامة والتي تطبق بشكل واسع في التحليل الاحصائي:

أولاً/ اختبار الاشارة للعينة الواحدة One Sample Sign Test

يستخدم هذا الاختبار عادةً للاستدلال على وسيط المجتمع خاصة عندما يكون التوزيع متماثل حيث يساوي الوسط الوسيط و يعتبر هذا الاختبار غير المعلمي بديلاً عن الاختبار المعلمي T - Test والذي نستخدمه في اختبار فرضيات حول وسط مجتمع واحد ويعتبر من أسهل الاختبارات غير المعلمية، والجدير بالذكر أن الكفاءة النسبية لهذا الاختبار بالنسب لاختبار T – Test تساوي 63.7%. أن البيانات التي يطبق عليها هذا الاختبار يجب أن تكون على الأقل من التدريج الترتيبي حيث الترتيب بالنسبة للوسيط المفترض، ويعتمد هذا الاختبار على عدد الاشارات الموجبة والتي تعبر عن اشارات الفروق بين قيم العينة وقيمة معينة نريد اختبار مساواتها لوسيط المجتمع وسندرس تطبيق الاختبار في الحالتين التاليتين.

- **اختبار فرضيات حول وسيط مجتمع واحد.**

بفرض أن لدينا مجتمع توزيعه متصل ووسيطه M ونريد اختبار الفرضية $\mathrm{H_0: M = M_0}$ مقابل احدى الفرضيات:

$$\mathrm{H_1: M \neq M_0}$$
$$\mathrm{H_1: M > M_0}$$
$$\mathrm{H_1: M < M_0}$$

حيث $\mathrm{M_0}$ هي قيمة معطاة، فإننا سنقوم بسحب عينة عشوائية $\{X_i\}_{i=1}^{n}$ ونحسب الفروق $X_i - \mathrm{M_0}$ ثم نحسب قيمة المعيار:

$$\boldsymbol{S = \#\, X_i : (X_i - \mathbf{M_0}) > 0}$$

حيث نستخدم الرمز # للدلالة على كلمة "عدد".

إن هذا المعيار يُعَبِّر عن متغير عشوائي بتوزيع ذي الحدين بغض النظر عن طبيعة المجتمع الذي لدينا، لذلك سنستخدم جداول توزيع ذي الحدين لتحديد القيم الحرجة ومناطق رفض وقبول الفرضية الصفرية. وقبل التطرق لإيجاد مناطق الرفض والقبول تذكر أن $\boldsymbol{p_r(X_i - \mathbf{M_0} = 0) = 0}$ وذلك لأن توزيع المجتمع هو توزيع متصل. وكذلك من تعريف الوسيط نستنتج أنه في حالة صحة الفرضية الصفرية فإن :

$$\boldsymbol{p_r(X_i - \mathbf{M_0} > 0) = p_r(X_i - \mathbf{M_0} < 0) = 0.5}$$

ويكون:

$$\boldsymbol{S \sim b(n, 0.5)}$$

وقبل حساب المعيار يتم حذف جميع قيم العينة التي تساوي M_0.

وسنميز هنا حالتين لتطبيق الاختبار وهما:

- **حجم العينة صغير $n \leq 20$**

إذا كانت الفرضية البديلة $H_1: M < M_0$ فإننا نحدد القيمة الحرجة من جداول توزيع ذي الحدين $b(n, 0.5)$ بحيث $p_r(S \leq s) \approx \alpha$. فمثلاً في حالة $n = 10$ و $\alpha = 0.5$ تكون $p_r(S \leq 2) \approx 0.055$ وعليه تكون 2 هي القيمة الحرجة وتكون منطقة الرفض Rejection Region هي $RR = \{0, 1, 2\}$ ، والجدير بالذكر هنا أننا قد لا نستطيع الحصول علي قيمة مستوى المعنوية المطلوب بالضبط من جدول ذي الحدين لذلك فننا نختار أقرب قيمة لها.

وإذا كانت الفرضية البديلة $H_1: M > M_0$ فإننا نجد القيمة الحرجة من جداول ذي الحدين بحيث $p_r(S \geq \grave{s}) \approx \alpha$ ، ولكن جداول ذي الحدين المرفقة في الكتاب تعطى احتمالات على الصورة $p_r(S \leq s)$ لذلك نستغل تماثل التوزيع $b(n, 0.5)$ وتكون $\grave{\boldsymbol{s}} = \boldsymbol{n} - \boldsymbol{s}$ ، فمثلاً عندما $n = 10$ و $\alpha = 0.5$ تكون $\grave{s} = 10 - 2 = 8$ وبالتالي تكون منطقة الرفض

$$RR = \{8, 9, 10\}$$

وإذا كانت الفرضية البديلة $H_1: M \neq M_0$ فإننا نجد قيمتين حرجتين s_2 , s_2 بحيث:

$$p_r(S \leq s_1) = p_r(S \geq s_2) \approx \frac{\alpha}{2}$$

حيث $s_2 = \boldsymbol{n} - s_1$.

مثال6.1: تمثل الأرقام التالية درجات الحرارة في مدينة في 15 يوماً من شهر مايو:

30,37,36,30,36,25,44,37,35,38,41,38,37,34,42

استخدم اختبار الاشارة للتحقق من الادعاء القائل بأن وسيط درجة الحرارة يساوي 39 وذلك عند مستوى معنوية 0.1

الحل/ لاحظ أنه لا يوجد أي مشاهدة في العينة تساوي 39

الفرضيات:

$$H_0: M = 39$$
$$H_1: M \neq 39$$

المنطقة الحرجة:

لاحظ أن التوزيع هنا هو $b(15,0.5)$ وتكون القيم الحرجة هي التي تحقق

$$p_r(S \leq s_1) = p_r(S \geq s_2) \approx 0.05$$

ومن جدول التوزيع نجد أن $p_r(S \leq 4) = 0.0592$ وبالتالي تكون القيم الحرجة هي :

$$s_1 = 4 \quad , \quad s_2 = 15 - 4 = 11$$

وتكون منطقة الرفض هي : $\mathrm{RR} = \{0,1,2,3,4,11,12,13,14,15\}$

المعيار:

X_i	42	34	37	38	41	38	35	37	44	25	36	30	36	37	30
sign	+	−	−	−	+	−	−	−	+	−	−	−	−	−	−

نلاحظ من الجدول أن :

$$S = \#\, X_i : (X_i - \mathrm{M}_0) > 0 = 3$$

المقارنة والقرار:

الطريقة الأولى: حيث أن $S = 3 \in \mathrm{RR}$ فإننا نرفض الفرضية الصفرية وبالتالي فإن وسيط درجة الحرارة لا يساوي 39

الطريقة الثانية: $\mathrm{p-value} = P(S \leq 3) = 0.0179 < \alpha = 0.05$ وبالتالي ترفض الفرضية الصفرية.

ملاحظة/ مستوى الدلالة الذي تم الاختبار عليه هو $0.1184 = 2 \times 0.0592$ وهو مختلف عن مستوى المعنوية المطلوب.

- **حجم العينة كبير $n > 20$**

ذكرنا سابقاً أنه في حالة صحة الفرضية الصفرية فإن $S \sim b(n, 0.5)$ حيث يكون توقع التوزيع $np = n \times 0.5 = \frac{n}{2}$ وتباين التوزيع $np(1-p) = \frac{n}{4}$ وبتطبيق نظرية النهاية المركزية يكون:

$$\boldsymbol{Z = \frac{S - \frac{n}{2}}{\frac{1}{2}\sqrt{n}} \sim N(0,1)}$$

في حالة الفرضية البديلة $\mathrm{H}_1 : \mathrm{M} < \mathrm{M}_0$ يكون الاختبار من اتجاه واحد نحو اليسار وتكون منطقة الرفض هي

$$\boldsymbol{\frac{S - \frac{n}{2}}{\frac{1}{2}\sqrt{n}} < -z_\alpha}$$

وبحل المتباينة نحصل على:

$$\boldsymbol{S < \frac{n}{2} - z_{\alpha} \times \frac{1}{2}\sqrt{n}}$$

و في حالة الفرضية البديلة $H_1: M > M_0$ يكون الاختبار من اتجاه واحد نحو اليمين و تكون منطقة الرفض هي

$$\boldsymbol{\frac{S - \frac{n}{2}}{\frac{1}{2}\sqrt{n}} > z_{\alpha}}$$

وبحل المتباينة نحصل على:

$$\boldsymbol{S > \frac{n}{2} + z_{\alpha} \times \frac{1}{2}\sqrt{n}}$$

و في حالة الفرضية البديلة $H_1: M \neq M_0$ يكون الاختبار من اتجاهين و تكون منطقة الرفض هي

$$\boldsymbol{\left(\frac{S - \frac{n}{2}}{\frac{1}{2}\sqrt{n}} < -z_{\frac{\alpha}{2}}\right) \cup \left(\frac{S - \frac{n}{2}}{\frac{1}{2}\sqrt{n}} > z_{\frac{\alpha}{2}}\right)}$$

وبحل المتباينة نحصل على:

$$\boldsymbol{\frac{n}{2} + z_{\frac{\alpha}{2}} \times \frac{\sqrt{n}}{2} < S < \frac{n}{2} - z_{\frac{\alpha}{2}} \times \frac{\sqrt{n}}{2}}$$

مثال6.2: اختار معلم 75 طالباً من تخصص المحاسبة الذين يدرسون مساق الاحصاء التطبيقي، وسألهم عن عدد الساعات التي يدرسونها أسبوعياً وكانت اجاباتهم كما هو موضح في الجدول التالي:

	أقل من 3 ساعات	3 ساعات	أكثر من 3 ساعات
عدد الطلبة	49	10	16

اختبر الادعاء بأن وسيط عدد ساعات الدراسة أقل من 3 ساعات عند مستوى دلالة 0.05

الحل/

سيتم استثناء العمود الثاني من حجم العينة وذلك لأن عدد الساعات يساوي 3 فتصبح $n = 65$

<u>الفرضيات:</u>

$$H_0: M = 3$$
$$H_1: M < 3$$

<u>المنطقة الحرجة:</u>

$$S < \frac{n}{2} - z_{0.05} \times \frac{1}{2}\sqrt{n}$$
$$S < \frac{65}{2} - 1.645 \times \frac{1}{2}\sqrt{65} \approx 26$$

<u>المعيار:</u>

من جدول البيانات المعطى نجد أن :

اقل من 3 ساعات اسبوعياً	~~3 ساعات اسبوعياً~~	اكثر من 3 ساعات اسبوعياً
49 طالب	~~10 طالب~~	16 طالب
−		+

وبالتالي $S = 16$

<u>المقارنة والقرار:</u>

حيث أن $S = 16 < 26$ فإننا نرفض الفرضية الصفرية وبالتالي فإن وسيط عدد ساعات الدراسة أقل من 3 ساعات.

- **اختبار الاشارة حول وسيطي مجتمعين مرتبطين.**

بفرض أن لدينا مجتمعين مرتبطين وسيطيهما M_1 و M_2 وكان المجتمعان يتبعان توزيعين متصلين ونريد اختبار الفرضية الصفرية $H_0: M_1 = M_2$ مقابل احدى الفرضيات:

$$H_1: M_1 \neq M_2$$
$$H_1: M_1 > M_2$$
$$H_1: M_1 < M_2$$

فإذا كانت X تمثل مشاهدة من المجتمع الاول و Y تمثل مشاهدة من المجتمع الثاني وكانت الفرضية الصفرية صحيحة فإن

$$P(X > Y) = P(X < Y)$$

وحيث أن التوزيعين متصلان فإن

$$P(X > Y) = P(X < Y) = \frac{1}{2}$$

وبالتالي يمكن كتابة الفرضية الصفرية على الشكل

$$\mathrm{H_0: p} = \frac{1}{2}$$

حيث $\mathrm{p} = P(X > Y)$ ويكون المعيار في هذه الحالة هو عدد المشاهدات التي يكون فيها $X > Y$ ويكون التوزيع المستخدم في الاختبار هو $b(n, 0.5)$ حيث n تمثل عدد المشاهدات المختلفة وذلك بعد حذف المشاهدات المتناظرة المتساوية في العينتين، ويتم اجراء الاختبار بنفس الآلية المتبعة في اختبار الاشارة لعينة واحدة.

مثال6.3: تقدم شخصان لوظيفة في شركة خاصة، فقررت إدارة الشركة تكليف 12 شخصاً من مجلس الإدارة لإجراء مقابلة مع الشخصين وكانت نتائج المقابلة كما يلي:

عضو لجنة المقابلة	درجة المتقدم الأول X	درجة المتقدم الثاني Y
1	10	8
2	7	5
3	5	7
4	6	6
5	8	7
6	8	5
7	7	7
8	6	3
9	6	7
10	5	3
11	7	5
12	8	6

هل تبين النتائج أفضلية لأحد المتقدمين على الآخر عند مستوى معنوية 0.01

الحل/ نفرض أن $p = P(X > Y)$، و لاحظ وجود مشاهدتين متساويتين وهما للعضو رقم 4 و العضو رقم 7 لذلك نعتبر $n = 10$

الفرضيات:

$$H_0: p = \frac{1}{2}$$
$$H_1: p \neq \frac{1}{2}$$

المنطقة الحرجة:

من جدول توزيع ذي الحدين نجد أن $P(S \leq 2) = 0.055$ وبالتالي تكون منطقة الرفض هي $RR = \{0,1,2,8,9,10\}$

المعيار:

حيث أن S هي عدد المشاهدات التي تكون فيها $X > Y$ فمن الجدول نجد أن $S = 8$

المقارنة والقرار:

حيث أن $S = 8 \in RR$ فإننا نرفض الفرضية الصفرية وبالتالي هناك أفضلية لأحد المتقدمين على الآخر.

مثال6.4: أعد حل مثال 6.3 باستخدام اختبار T للفروقات بين مشاهدات بفرض أن مجتمع الفروق طبيعي.

الحل/

باعتبار أن $D_i = X_i - Y_i$ نجد أن:

$$\mu_d = \frac{\sum D_i}{n} = \frac{14}{12} = 1.167$$
$$S_D^2 = \frac{\sum (D_i - \mu_d)^2}{n-1} = \frac{27.67}{11} = 2.52$$

الفرضيات:

$$H_0: \mu_D = 0$$
$$H_1: \mu_D \neq 0$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع T وذلك لصغر حجم العينة وحيث أن $\alpha = 0.1$ والاختبار ذو اتجاهين إذن تكون القيم الحرجة:

$$t_{(0.05,11)} = 1.796 \quad \text{،} \quad -t_{(0.05,11)} = -1.796$$

المعيار:

$$\mathrm{T} = \frac{\mu_d}{S_d/\sqrt{n}} = \frac{1.167}{1.59/\sqrt{11}} = 2.43$$

المقارنة والقرار:

وحيث أن $(t_{(0.05,11)} = 1.796) < (\mathrm{T} = 2.43)$ فإننا نرفض الفرضية الصفرية وعليه نعتبر أنه يوجد أفضلية لأحد المتقدمين على الآخر.

مثال6.5: في استطلاع لرأي الطلبة حول تفضيلهم لنوعين من المشروب، تم استطلاع رأي 110 طلاب، وكانت نتائج الاستطلاع كما يلي:

	يفضلون الشاي	يفضلون القهوة	لا فرق بين المشروبين
عدد الطلبة	37	63	10

هل تعطي نتيجة الاستطلاع دليلاً كافياً على وجود فرق بين النوعين من حيث الافضلية عند مستوى معنوية 0.05

الحل/ نفرض أن **p** تمثل أفضلية الشاي، ولاحظ أن $n = 100$ حيث تم استثناء 10 أشخاص تتساوي الأفضلية عندهم للمشروبين.

الفرضيات:

$$\mathrm{H_0: p} = \frac{1}{2}$$

$$\mathrm{H_1: p} \neq \frac{1}{2}$$

المنطقة الحرجة:

حيث أن حجم العينة $n = 100 > 20$ فإننا سنقرب التوزيع للتوزيع الطبيعي وبالتالي تكون المنطقة الحرجة كما يلي:

$$\frac{n}{2} + z_{\frac{\alpha}{2}} \times \frac{\sqrt{n}}{2} < S < \frac{n}{2} - z_{\frac{\alpha}{2}} \times \frac{\sqrt{n}}{2}$$

$$\frac{100}{2} + 1.96 \times \frac{\sqrt{100}}{2} < S < \frac{100}{2} - 1.96 \times \frac{\sqrt{100}}{2}$$

$$59.8 < S < 40.2$$

وبصورة أخرى تكون منطقة الرفض $\mathrm{RR} = (-\infty, 40.2) \cup (59.8, \infty)$

المعيار:

حيث أن S هي عدد المشاهدات التي تكون فيها الأفضلية للشاي فيكون $S = 37$

المقارنة والقرار:

حيث أن $S = 37 \in \mathrm{RR}$ فإننا نرفض الفرضية الصفرية وبالتالي هناك أفضلية لأحد المشروبين على الآخر.

ثانياً/ اختبار اشارة الرتب (ولكيكسون) Wilcoxon Signed-Rank Test

يُطبق هذا الاختبار غير المعلمي عند اختبار تساوي وسيطي مجتمعين مرتبطين توفر لدينا معلومات رقمية – تتدرج ضمن تدريج فترة أو نسبة – عن عينتين منهما $\{X_i\}_{i=1}^{n}$ ، $\{Y_i\}_{i=1}^{n}$ حيث ترتبط كل X_i بالقيمة المناظرة لها Y_i وهناك ما يمنع من تطبيق الاختبارات المعلمية، و في هذه الحالة تطبيق اختبار الاشارة هنا يفقدنا جزءاً مهماً من البيانات المتوفرة، وكفاءة هذا الاختبار بالنسبة لاختبار T – Test تساوي 95% وبالتالي يعتبر أفضل من اختبار الاشارة السابق الذكر.

وتكون الفرضية الصفرية $\mathrm{H_0: M_1 = M_2}$ مقابل احدى الفرضيات:

$$\mathrm{H_1: M_1 \neq M_2}$$
$$\mathrm{H_1: M_1 > M_2}$$
$$\mathrm{H_1: M_1 < M_2}$$

ويتم تحديد مناطق الرفض من جدول ولكيكسون حسب الفرضية البديلة ومستوى المعنوية المستخدم في الاختبار، ففي حالة الفرضية البديلة $\boldsymbol{H_1: M_1 > M_2}$ يكون المعيار المستخدم هو $\boldsymbol{T^+}$ والذي يعبر عن مجموع الرتب المقابلة للفروق الموجبة و نختار القيمة الحرجة $\boldsymbol{x}$ من الجدول وهي القيمة التي تحقق $\boldsymbol{P(T^+ \geq x) \approx \alpha}$.

وفي حالة الفرضية $\boldsymbol{H_1: M_1 < M_2}$ يكون المعيار المستخدم هو $\boldsymbol{T^-}$ والذي يعبر عن مجموع الرتب المقابلة للفروق السالبة ونختار أيضاً القيمة التي تحقق $\boldsymbol{P(X \geq x) \approx \alpha}$.

أما في حالة الفرضية $H_1: M_1 \neq M_2$ فنختار المعيار حسب الاشارات الأكثر تعداد ونستخدم الاختبار من جهة اليمين. وسنميز هنا حالتين:

- **حجم العينة $n \leq 15$**

يتم تطبيق الاختبار حسب الخطوات التالية:

- نحسب الفروق بين المشاهدات المتناظرة D_i.
- نوجد القيم المطلقة لهذه الفروق $|D_i|$.
- نحسب رتب القيم المطلقة للفروق وسنفرض هنا عدم وجود رتب متساوية.
- حدد إشارة الفروق المناظرة لكل رتبة.
- حدد منطقة رفض الفرضية الصفرية من جدول ولكيكسون.
- نحسب مجموع الرتب المناظرة للفروق الموجبة أو السالبة حسب الفرضية البديلة.
- قرر قبول أو رفض الفرضية الصفرية حسب وقوع المعيار في منطقة القبول أو الرفض.

مثال6.6: الجدول التالي يبين عدد الأميال التي تقطعها 12 سيارة باستخدام نوعين من الوقود.

الوقود A	26.4	10.3	15.8	16.5	32.5	8.3	22.1	30.1	12.9	12.6	27.3	9.4
الوقود B	24.3	9.8	16.9	17.2	30.5	7.9	22.4	28.6	13.1	11.6	25.5	8.6

اختبر الادعاء بأن وسيط A اكبر من وسيط B عند مستوى معنوية 0.05

الحل/

نكون الجدول التالي حيث $D_i = A_i - B_i$:

A	26.4	10.3	15.8	16.5	32.5	8.3	22.1	30.1	12.9	12.6	27.3	9.4
B	24.3	9.8	16.9	17.2	30.5	7.9	22.4	28.6	13.1	11.6	25.5	8.6
D_i	2.1	0.5	−1.1	−0.7	2	0.4	−0.3	1.5	−0.2	1	1.8	0.8
$\|D_i\|$	2.1	0.5	1.1	0.7	2	0.4	0.3	1.5	0.2	1	1.8	0.8
رتب $\|D_i\|$	12	4	8	5	11	3	2	9	1	7	10	6
Sign	+	+	−	−	+	+	−	+	−	+	+	+

الفرضيات:

$$H_0: M_A = M_B$$
$$H_1: M_A > M_B$$

المنطقة الحرجة:

من جدول توزيع $\mathbf{T}^+$ بحجم عينة $n = 12$ نجد أن:

$$P(T^+ \geq 61) = 0.046$$

وهي أقرب قيمة لمستوى المعنوية المطلوب، وبالتالي تكون منطقة الرفض $RR = [61, \infty)$

المعيار:

$$T^+ = 12+4+11+3+9+7+10 = 62$$

المقارنة والقرار:

حيث أن $T^+ = 62 \in RR$ فإننا نرفض الفرضية الصفرية وبالتالي نستطيع القول بأن وسيط A أكبر من وسيط B.

- **حجم العينة $n > 15$**

في هذه الحالة يتم تقريب توزيع T^+ للتوزيع الطبيعي حيث:

$$\boldsymbol{E}(T^+) = \frac{n(n+1)}{4}$$

$$\boldsymbol{\sigma}_{T^+} = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

ويكون المعيار

$$\boldsymbol{Z} = \frac{\mathbf{T}^+ - \frac{\boldsymbol{n}(\boldsymbol{n}+\mathbf{1})}{\mathbf{4}}}{\sqrt{\frac{\boldsymbol{n}(\boldsymbol{n}+\mathbf{1})(\mathbf{2}\boldsymbol{n}+\mathbf{1})}{\mathbf{24}}}} \sim \boldsymbol{N}(\mathbf{0}, \mathbf{1})$$

وذلك في حالة عدم وجود تساوي في الرتب المطلقة، وإذا وُجِدَ تساوي في الرتب المطلقة فيتم حساب المعيار كالتالي:

بفرض أن:

$\grave{n}$: تعبر عن عدد الفروق غير الصفرية.

k: عدد حالات الحصول على رتب متساوية للفروق المطلقة.

q_i: عدد الفروق التي لها نفس الرتب في الحالة i.

فيكون التوقع والتباين و المعيار كما يلي:

$$\boldsymbol{E}(\mathrm{T}^{+}) = \frac{\grave{n}(\grave{n}+1)}{4}$$

$$\boldsymbol{\sigma}_{\mathrm{T}^{+}} = \sqrt{\frac{\grave{n}(\grave{n}+1)(2\grave{n}+1)}{24} - \frac{1}{48}\sum_{i=1}^{k} q_i({q_i}^2-1)}$$

ويكون المعيار

$$\boldsymbol{Z} = \frac{\mathbf{T}^{+} - \dfrac{\boldsymbol{\grave{n}}(\boldsymbol{\grave{n}}+\mathbf{1})}{\mathbf{4}}}{\boldsymbol{\sigma}_{\mathrm{T}^{+}}} \sim \boldsymbol{N}(\mathbf{0},\mathbf{1})$$

مثال6.7: الجدول التالي يبين مقاييس تفضيل 20 متطوع على نوعين من مكونات وجبة الافطار.

الوجبة A	70	85	73	75	65	50	80	71	80	51
	72	76	79	65	59	72	84	90	56	57
الوجبة B	65	41	45	80	84	50	71	52	42	78
	62	38	80	65	54	67	87	90	38	14

اختبر عند مستوى معنوية 0.05 وجود فروق بين وسيط A ووسيط B

الحل/

نكون الجدول التالي:

sign	رتب $\lvert D_i \rvert$	$\lvert D_i \rvert$	D_i	B	A
+	4.5	5	5	65	70
+	17	44	44	41	85
+	13	28	28	45	73
−	4.5	5	−5	80	75
−	10.5	19	−19	84	65
		0	0	50	50
+	7	9	9	71	80
+	10.5	19	19	52	71
+	14.5	38	38	42	80
−	12	27	−27	78	51
+	8	10	10	62	72
+	14.5	38	38	38	76
−	1	1	−1	80	79
		0	0	65	65
+	4.5	5	5	54	59
+	4.5	5	5	67	72
−	2	3	−3	87	84
		0	0	90	90
+	9	18	18	38	56
+	16	43	43	14	57

لاحظ هنا ما يلي:

$$T^{+} = 4.5 + 4.5 + 4.5 + 7 + 8 + 9 + 10 + 11.5 + 14 + 15.5$$
$$+15.5 + 17 = 121$$

وعدد الفروق غير الصفرية $\grave{n} = 17$ ، وعدد حالات الحصول على رتب متساوية $\mathrm{k} = 3$ وعدد الفروق التي لها نفس الرتب هي:

$q_1 = 4$ للرتبة 4.5 ، $q_2 = 2$ للرتبة 10.5 ، $q_3 = 2$ للرتبة 14.5

وبالتالي:

$$E(\mathrm{T}^{+}) = \frac{17 \times 18}{4} = 76.5$$

$$\sigma_{\mathrm{T}^{+}} = \sqrt{\frac{17 \times 18 \times 35}{24} - \frac{1}{48}[4 \times 15 + 2 \times 3 + 2 \times 3]} = 21.1$$

<u>الفرضيات:</u>

$$\mathrm{H}_0: \mathrm{M}_A = \mathrm{M}_B$$
$$\mathrm{H}_1: \mathrm{M}_A \neq \mathrm{M}_B$$

<u>المنطقة الحرجة:</u>

حيث أن الاختبار من اتجاهين نجد من جدول توزيع $Z_{\frac{\alpha}{2}} = 1.96$ أن $\boldsymbol{Z}$ و $-Z_{\frac{\alpha}{2}} = -1.96$

<u>المعيار:</u>

$$Z = \frac{121 - 76.5}{21.1} = 2.11$$

<u>المقارنة والقرار:</u>

حيث أن $(Z_{\frac{\alpha}{2}} = 1.96) < (Z = 2.11)$ فإننا نرفض الفرضية الصفرية وبالتالي نستطيع القول بأنه يوجد فروق معنوية بين الوسيطين.

ثالثاً/ اختبار مان– وتني Mann-Whitney Test

يعرف هذا الاختبار باختبار **U** (**U Test**) ويعتبر بديلاً عن اختبار T المعلمي الذي يختبر الفرق بين متوسطي مجتمعين طبيعيين مستقلين، حيث في حالة تعذر تطبيق الاختبار المعلمي يتم تطبيق اختبار مان– وتني والذي يستخدم رتب المشاهدات لاختبار وجود فرق بين وسيطي المجتمعين وهذا الاختبار ذو اتجاه واحد نحو اليسار وسنميز هنا حالتين:

- **في حالة $n_1, n_2 \leq 10$**

نفرض أن لدينا مجتمعين مستقلين يتبعان توزيعين متصلين وأخذنا عينة عشوائية $\{X_i\}_{i=1}^{n_1}$ من المجتمع الاول وعينة عشوائية $\{Y_i\}_{i=1}^{n_2}$ ونريد اختبار الفرضية الصفرية $\mathrm{H_0: M_1 = M_2}$ مقابل احدى الفرضيات:

$$\mathrm{H_1: M_1 \neq M_2}$$
$$\mathrm{H_1: M_1 > M_2}$$
$$\mathrm{H_1: M_1 < M_2}$$

فإننا نقوم بالخطوات التالية:

- ترتيب مشاهدات العينة المشتركة والتي حجمها $n_1 + n_2$ تصاعدياً، وفي حال تكرار قيمة مشاهدة معينة في العينة المشتركة تمنح جميعها نفس الرتبة وهي الوسط الحسابي لرتب هذه المشاهدات.
- نجمع رتب كل عينة على حده حيث نرمز لمجموع رتب العينة $\{X_i\}_{i=1}^{n_1}$ بالرمز $\mathrm{T_x}$ ولمجموع رتب العينة $\{Y_i\}_{i=1}^{n_2}$ بالرمز $\mathrm{T_y}$.
- نحسب القيمتين التاليتين:
-

$$\mathbf{U}_x = \boldsymbol{n_1 n_2} + \frac{\boldsymbol{n_1(n_1+1)}}{\mathbf{2}} - \mathbf{T}_x$$
$$\mathbf{U}_y = \boldsymbol{n_1 n_2} + \frac{\boldsymbol{n_2(n_2+1)}}{\mathbf{2}} - \mathbf{T_y}$$

- نجد منطقة الرفض باستخدام جدول U وهي $(-\infty, \mathrm{U}_\mathbf{0}]$ ، حيث في حالة الفرضية $\mathrm{H_1: M_1 \neq M_2}$ تحسب $\mathrm{U}_\mathbf{0}$ من العلاقة $P(U \leq \mathrm{U}_\mathbf{0}) = \frac{\alpha}{2}$، وفي الحالتين الاخريين تحسب $\mathrm{U}_\mathbf{0}$ من العلاقة $P(U \leq \mathrm{U}_\mathbf{0}) = \alpha$

- نحسب المعيار وهو:

$$\mathbf{U} = \mathbf{min}(\mathbf{U}_x, \mathbf{U}_y)$$

- نقرر قبول الفرضية الصفرية أو رفضها حسب وقوع قيمة المعيار في منطقة القبول أو الرفض.

مثال6.8: قرر معلم اجراء اختبار لمجموعة من الطلبة بعمل نموذجين للاختبار واعطاء نصف الطلبة نموذج معين والنصف الثاني النموذج الثاني، وبعد انجاز الاختبار وتقدير الدرجات رصد المعلم الدرجات في الجدول التالي:

المجموعة الأولى	المجموعة الثانية
52	72
78	62
56	91
90	88
65	90
86	74
64	98
90	80
49	81
78	71

اختبر إذا ما كان النموذجان متكافئان أم لا وذلك عند مستوى معنوية 0.1

الحل/

نكون جدول الرتب وهو كما يلي:

رتب Y	رتب X
8	2
4	10.5
19	3
15	17
17	6
9	14
20	5
12	17
13	1
7	10.5
$\mathbf{T_y = 124}$	$\mathbf{T_x = 86}$

وحيث أن $n_1 = n_2 = 10$ إذن:

$$U_x = n_1 n_2 + \frac{n_1(n_1+1)}{2} - T_x$$

$$U_x = 10 \times 10 + \frac{10(10+1)}{2} - 86 = 69$$

$$U_y = n_1 n_2 + \frac{n_2(n_2+1)}{2} - T_y$$

$$U_y = 10 \times 10 + \frac{10(10+1)}{2} - 124 = 31$$

<u>الفرضيات:</u>

$$H_0: M_1 = M_2$$
$$H_1: M_1 \neq M_2$$

<u>المنطقة الحرجة:</u>

حيث أن الاختبار ذو اتجاهين وبالتالي تكون $P(U \leq U_0) = 0.05$ ومن جدول U نحصل على $U_0 = 23$ وبالتالي تكون منطقة الرفض $RR = (-\infty, 23]$

المعيار:

$$U = \min(69,31) = 31$$

المقارنة والقرار:

حيث أن $U = 31 \notin RR$ فإننا نقبل الفرضية الصفرية وبالتالي لا يوجد فروق ذات دلالة احصائية بين الاختبارين.

- **في حالة $n_1, n_2 > 10$**

في هذه الحالة يتم تقريب توزيع U لتوزيع ذات الحدين حيث:

$$E(U) = \frac{n_1 n_2}{2}$$

$$\sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

وبالتالي يكون:

$$\frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} \sim N(0,1)$$

فإذا كان الاختبار ذو اتجاهين فإن:

$$P(U \leq U_0) = \frac{\alpha}{2}$$

ومنها يمكن استنتاج أن:

$$U_0 = \frac{n_1 n_2}{2} - Z_{\frac{\alpha}{2}} \times \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

وإذا كان الاختبار ذو اتجاه واحد فإن:

$$P(U \leq U_0) = \alpha$$

ومنها يمكن استنتاج أن:

$$U_0 = \frac{n_1 n_2}{2} - Z_{\alpha} \times \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

مثال6.9: قام مهندس زراعي بتسميد 18 شتلة من أشتال البندورة بالسماد الكيماوي وتسميد 15 شتلة أخرى بالسماد الطبيعي، وعند الحصاد تم وزن انتاج كل شتلة بالكجم فكانت كما يلي:

السماد الكيماوي X	1.25	1.35	1.57	1.31	1.39	1.47	1.54	1.05	1.53
	1.56	1.11	1.08	1.49	1.49	1.51	1.33	1.56	1.58
السماد الطبيعي Y	1.19	1.45	1.08	1.35	1.32	1.24	1.63	1.59	1.33
	1.37	1.59	1.42	1.36	1.47	1.51			

هل تستطيع أن تستنتج أن السماد الكيماوي يعطي محصولاً أكثر من السماد الطبيعي عند مستوى معنوية 0.05

الحل/ لاحظ أن $n_1 = 18$, $n_2 = 15$

رتب X		رتب Y	
7	27.5	5	15
12.5	4	19	31.5
29	2.5	2.5	18
8	17	12.5	14
16	22	9	20.5
20.5	23.5	6	23.5
26	10.5	33	
1	27.5	31.5	
25	30	10.5	
$\mathbf{T_x = 309.5}$		$\mathbf{T_y = 251.5}$	

$$U_x = n_1 n_2 + \frac{n_1(n_1+1)}{2} - T_x$$

$$U_x = 18 \times 15 + \frac{18(18+1)}{2} - 309.5 = 131.5$$

$$U_y = n_1 n_2 + \frac{n_2(n_2+1)}{2} - T_y$$

$$U_y = 18 \times 15 + \frac{15(15+1)}{2} - 251.5 = 138.5$$

الفرضيات:

$$H_0: M_X = M_Y$$
$$H_1: M_X > M_Y$$

المنطقة الحرجة:

حيث أن $n_1, n_2 > 10$ سنقرب التوزيع للتوزيع الطبيعي ولاحظ أن الاختبار ذو اتجاه واحد لذلك حيث: $P(U \leq U_{\mathbf{0}}) = 0.05$

$$U_0 = \frac{n_1 n_2}{2} - Z_\alpha \times \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

$$U_0 = \frac{18 \times 15}{2} - 1.645 \times \sqrt{\frac{18 \times 15(18 + 15 + 1)}{12}} = 90$$

وبالتالي تكون منطقة الرفض $RR = (-\infty, 90]$

المعيار:

$$U = \min(131.5, 138.5) = 131.5$$

المقارنة والقرار:

حيث أن $U = 131.5 \notin RR$ فإننا نقبل الفرضية الصفرية وبالتالي لا يوجد دليل على أن السماد الكيماوي يعطي انتاجاً أعلى من السماد الطبيعي.

رابعاً/ اختبار كروسكال– والاس Kruskal – Wallace Test

يعتبر هذا الاختبار تعميماً لاختبار مان – وتني للمتغيرات المتصلة ويعتبر بديلاً لتحليل التباين الأحادي في الاختبارات المعلمية، حيث أنه يستخدم لمقارنة عدة أوساط لمجتمعات مستقله، وهو اختبار ذو اتجاه واحد نحو اليمين. أن الآلية المتبعة في حساب معيار الاختبار هي نفسها المتبعة في اختبار مان وتني. فمثلاً إذا كان لدينا k معالجة وعدد الاستجابات في كل عينة هو n_i حيث $i = 1,2, \dots , k$ ونظمت البيانات في الجدول التالي:

المعالجات	1		j		n_i
1	y_{11}	..	y_{1j}	..	y_{1n_i}
.	.		.		.
.	.		.		.
i	y_{i1}	..	y_{ij}	..	y_{in_i}
.	.		.		.
.	.		.		.
k	y_{k1}	..	y_{kj}	..	y_{kn_i}

يتم ترتيب جميع البيانات تصاعدياً مع اعطاء الاستجابات المكررة رتباً تساوي الوسط الحسابي لأرقام ترتيبها المتسلسلة، ثم نجمع رتب كل معالجة وبفرض أن هذه الرتب ومجاميعها كما هو موضح بالجدول التالي:

المعالجات	1		j		n_i	مجموع رتب المعالجة
1	r_{11}	..	r_{1j}	..	r_{1n_i}	R_1
.	.		.		.	.
.	.		.		.	.
i	r_{i1}	..	r_{ij}	..	r_{in_i}	R_i
.	.		.		.	.
.	.		.		.	.
k	r_{k1}	..	r_{kj}	..	r_{kn_i}	$\mathrm{R_k}$

ويكون المعيار المستخدم هو:

$$\mathbf{H} = \frac{\mathbf{12}}{\boldsymbol{n}(\boldsymbol{n}+\mathbf{1})}\left[\frac{\mathbf{R}_1^2}{\boldsymbol{n}_1} + \frac{\mathbf{R}_2^2}{\boldsymbol{n}_2} + \cdots + \frac{\mathbf{R}_k^2}{\boldsymbol{n}_k}\right] - \mathbf{3}(\boldsymbol{n}+\mathbf{1})$$

ومن الجدير بالذكر هنا أن هذا المعيار يكون قريباً جداً من توزيع $\chi^2_{(k-1)}$ بمعنى:

$$\mathrm{H} \sim \chi^2_{(k-1)}$$

مثال6.10: لمقارنة ثلاثة أنواع من الأدوية لمعالجة الصداع، أخذت مجموعة من 22 شخصاً يعانون من الصداع، وقسموا إلى ثلاث مجموعات، وأُعطيت كل مجموعة نوعاً من الأدوية وتم رصد زمن الشفاء بالدقائق وكانت النتائج كما يلي:

T_1	T_2	T_3
80	56	58
53	22	52
55	44	41
56	46	53
65	29	35
56	34	21
70	38	54
		47

اختبر الفرضية القائلة بأنه لا يوجد فروق بين الأدوية الثلاثة عند معنوية 0.05

الحل/

لاحظ هنا أن $n_1 = 7\ ,\ n_2 = 7\ ,\ n_3 = 8$ وبالتالي $n = 22$ ، نحسب الآن الرتب ومجاميعها للمعالجات الثلاثة كما هو موضح بالجدول التالي:

	T_1	T_2	T_3
	22	17	19
	12.5	2	11
	15	8	7
	17	9	12.5
	20	3	5
	17	4	1
	21	6	14
			10
مجموع رتب المعالجات	$R_1 = 124.5$	$R_2 = 49.5$	$R_3 = 79.5$

<u>الفرضيات:</u>

H_0: لا يوجد فروق بين الأدوية

H_1: يوجد فروق بين الأدوية

<u>المنطقة الحرجة:</u>

حيث أن التوزيع هو توزيع كاي تربيع بدرجة حرية 2 و حيث أن مستوى المعنوية 0.05 فيكون:

$$\chi^2_{(0.05,2)} = 5.991$$

<u>المعيار:</u>

$$H = \frac{12}{n(n+1)}\left[\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \cdots + \frac{R_k^2}{n_k}\right] - 3(n+1)$$

$$H = \frac{12}{22 \times 23}\left[\frac{(124.5)^2}{7} + \frac{(49.5)^2}{7} + \frac{(79.5)^2}{8}\right] - 3 \times 23 = 10.38$$

<u>المقارنة والقرار:</u>

حيث أن $(\chi^2_{(0.05,2)} = 5.991) > H = 10.38$ فإننا نرفض الفرضية الصفرية وبالتالي يوجد دليل على أنه يوجد اختلاف بين الأدوية الثلاثة عند مستوى معنوية 0.05.

خامساً/ اختبارات كاي تربيع للاستقلالية وجودة المطابقة غير المعلمية

- **اختبار الاستقلالية Independence Test**

يستخدم هذا الاختبار عند رغبة باحث معرفة إذا ما كان متغيران مستقلان أم لا وهناك مانع من تطبيق اختبار الاستقلالية المعلمي (اختبار معامل الارتباط) ، فمثلاً نستخدم هذا الاختبار عند الرغبة في معرفة تأثير لقاح معين في الاصابة بمرض معين أو دراسة ما إذا كان النجاح و الرسوب في مساق معين يعتمد على مدرس المساق أم لا وهكذا.

بفرض أن A , B ظاهرتين معينتين ونريد أن نختبر الفرضية H_0: A و B مستقلان مقابل H_1: A و B غيرمستقلان وكانت كل خاصية بعدد معين من المستويات، فمثلاً الخاصية A مستوياتها $A_1, A_2, \dots, A_r$ والخاصية B على مستويات $B_1, B_2, \dots, B_k$ يكون لدينا عدد $r \times k$ من الخلايا، نأخذ عينة عشوائية من مجتمع الدراسة ونعد عدد الاستجابات الفعلي في كل خلية وسنرمز له بالرمز O_{ij} والذي يدل على عدد الاستجابات في الخلية A_iB_j ونسجلها في جدول كما يلي:

A \ B	B_1	.	B_j	.	B_k	المجموع
A_1	O_{11}	.	O_{1j}	.	O_{1k}	R_1
.	.	.	.	.	.	
A_i	O_{i1}	.	O_{ij}	.	O_{ik}	R_i
.	.	.	.	.	.	
A_r	O_{r1}	.	O_{rj}	.	O_{rk}	R_r
المجموع	C_1		C_j		C_k	n

إذا كانت الفرضية الصفرية صحيحة فإن احتمال وقوع مشاهدة في الخلية A_iB_j هو:

$$p_{ij}\{A_iB_j\} = p_{ij}\{A_i\}p_{ij}\{B_j\} = \frac{R_i}{n} \times \frac{C_j}{n}$$

وبناءً عليه يكون عدد الاستجابات المتوقع في الخلية A_iB_j وسنرمز له بالرمز $\boldsymbol{e_{ij}}$ هو:

$$\boldsymbol{e_{ij}} = n \times p_{ij}\{A_iB_j\} = n \times \frac{R_i}{n} \times \frac{C_j}{n} = \frac{R_iC_j}{n}$$

وبالتالي تكون الفروقات $O_{ij} - e_{ij}$ يمكن أن تؤخذ كمقياس لصدق الفرضية الصفرية أو عدم صدقها، حيث الفروق الصغيرة تدعم الفرضية الصفرية والفروق الكبيرة تدعم عدم صحتها ومعيار الاختبار المبني على هذه الفروق يعطى بالصيغة التالية:

$$\chi^2 = \sum_{j=1}^{k} \sum_{i=1}^{r} \frac{(e_{ij} - O_{ij})^2}{e_{ij}}$$

وهذا المعيار قريب جداً من توزيع كاي تربيع بدرجة حرية $(r-1)(k-1)$ والاختبار هنا ذو اتجاه واحد نحو اليمين ويوصى دائماً أن لا تقل تكرارات التجربة عن 50 وتكرارات كل خلية عن 5 مشاهدات وفي حال وجود خلية تكرارها أقل من 5 تدمج مع الخلية التي تسبقها أو تليها.

مثال6.11: أُجريت تجربة لتقييم فعالية تطعيم معين ضد الرشح حيث شملت التجربة 1000 شخص، أخذت منهم مجموعة جرعة واحدة من التطعيم، ومجموعة أخرى أخذت جرعتين، أما المجموعة الأخيرة فلم تأخذ التطعيم. وتمت ملاحظة الأشخاص من حيث إصابتهم بالرشح أم لا فكانت النتائج كما يلي:

	جرعة واحدة	جرعتين	بدون تطعيم	المجموع
أصيبوا بالرشح	24	9	13	46
لم يصابوا	289	100	565	954
المجموع	313	109	578	1000

هل تعطي النتائج دليلاً كافياً على أن التطعيم يحد من الإصابة بالرشح عند معنوية 0.05 ؟

الحل/

نكون جدول التوافق التالي:

جدول التوافق

O_{ij}	e_{ij}	$e_{ij}-O_{ij}$	$(e_{ij}-O_{ij})^2$	$\frac{(e_{ij}-O_{ij})^2}{e_{ij}}$
24	14.4	9.6	92.16	6.4
9	5	4	16	3.2
13	26.6	−13.6	184.96	6.95
289	298.6	−9.6	92.16	0.31
100	104	−4	16	0.15
565	551.4	13.6	184.96	0.34
				17.35

الفرضيات:

H_0:التطعيم لا يحد من الإصابة بالرشح

H_1:التطعيم يحد من الإصابة بالرشح

المنطقة الحرجة:

حيث أن التوزيع هو توزيع كاي تربيع بدرجة حرية 2=(1−2)(1−3) و حيث أن مستوى المعنوية 0.05 والاختبار ذو اتجاه واحد نحو اليمين فيكون:

$$\chi^2_{(0.05,2)} = 5.991$$

المعيار:

$$\chi^2 = \sum_{j=1}^{k}\sum_{i=1}^{r}\frac{(e_{ij}-O_{ij})^2}{e_{ij}} = 17.35$$

المقارنة والقرار:

حيث أن $\chi^2 = 17.35 > (\chi^2_{(0.05,2)} = 5.991)$ فإننا نرفض الفرضية الصفرية وبالتالي يوجد دليل على أن التطعيم يحد من الرشح عند مستوى معنوية 0.05.

- **اختبار التجانس Homogeneity Test**

يستخدم هذا الاختبار لاختبار وجود تساوي نسبة صفة معينة في عدة مجتمعات، فبفرض أن لدينا مجتمعاً ما وقمنا بتقسيمه لعدة مجتمعات جزئية، كأن يكون المجتمع جامعة ما والمجتمعات الجزئية هي الكليات، وأخذنا عينة من كل هذه المجتمعات الجزئية وكان الهدف هو اختبار وجود صفة معينة بنفس النسبة في هذه المجتمعات أم هناك فروق في هذه النسب، نستخدم في هذه الحالة اختبار التجانس، وهو شبيه باختبار الاستقلالية ويختلف عنه في أن حجوم العينات محدد مسبقاً بينما في الاستقلال يكون الحجم الكلي فقط هو المحدد سلفاً وحجوم الصفوف والاعمدة متغيرات عشوائية وهناك فرق آخر وهو أننا في اختبار التجانس نقيس تساوي نسب خاصية في المجتمعات أما في اختبار الاستقلالية فنختبر استقلالية عاملين.

إذا كان لدينا عدد r من المجتمعات المستقلة هي $\mathrm{A}_1, \mathrm{A}_2, \dots, \mathrm{A}_\mathrm{r}$ وسحبت عينات من المجتمعات عددها k هي $\mathrm{B}_1, \mathrm{B}_2, \dots, \mathrm{B}_k$ وتم عد الاستجابات الموجودة في كل مجتمع من كل عينة وكانت النتائج كما يلي:

العينات / المجتمعات	B_1	.	B_j	.	B_k	المجموع
A_1	O_{11}	.	O_{1j}	.	O_{1k}	R_1
.	.	.	.	.	.	
A_i	O_{i1}	.	O_{ij}	.	O_{ik}	R_i
.	.	.	.	.	.	
A_r	O_{r1}	.	O_{rj}	.	O_{rk}	R_r
المجموع	C_1		C_j		C_k	n

نكون جدول توافق مشابه لجدول اختبار الاستقلالي ونستخدم المعيار التالي:

$$\chi^2 = \sum_{j=1}^{k}\sum_{i=1}^{r}\frac{(e_{ij} - O_{ij})^2}{e_{ij}}$$

وهذا المعيار قريب جداً من توزيع كاي تربيع بدرجة حرية $(r-1)(k-1)$ والاختبار هنا ذو اتجاه واحد نحو اليمين.

مثال6.12: تم تقسيم أحد المجتمعات إلى أربعة مجتمعات جزئية هي العمال، المدرسين، الطلبة والتجار وتم سؤالهم حول مدى رضاهم عن سلوك المشاة داخل المدن المكتظة. وقد شملت الدراسة على 300 من العمال و 250 من المدرسين و 300 طالب و 350 تاجر وكانت النتائج كما يلي:

	راضي	غير راضي	حجم العينة
العمال	32	268	300
المدرسين	51	199	250
الطلبة	67	233	300
التجار	83	267	350
المجموع	233	967	1200

اختبر فيما إذا كانت نسبة الرضى عن سلوك المشاة متساوية وفق آراء الفئات الأربعة عند مستوى معنوية 0.05

الحل/

نكون جدول التوافق التالي:

جدول التوافق

O_{ij}	e_{ij}	$e_{ij}-O_{ij}$	$(e_{ij}-O_{ij})^2$	$\frac{(e_{ij}-O_{ij})^2}{e_{ij}}$
32	58.25	−26.25	689.0625	11.83
51	48.54	2.46	6.0516	0.12
67	58.25	8.75	76.5625	1.31
83	67.96	15.04	226.2016	3.33
268	241.75	26.25	689.0625	2.85
199	201.46	−2.46	6.0516	0.03
233	241.75	−8.75	76.5625	0.32
267	282.04	−15.04	226.2016	0.80
				20.59

الفرضيات:

H_0: نسبة الرضى متساوية في جميع الفئات

H_1: نسبة الرضى غير متساوية في جميع الفئات

المنطقة الحرجة:

حيث أن التوزيع هو توزيع كاي تربيع بدرجة حرية $3=(4-1)(2-1)$ و حيث أن مستوى المعنوية 0.05 والاختبار ذو اتجاه واحد نحو اليمين فيكون:

$$\chi^2_{(0.05,3)} = 7.815$$

المعيار:

$$\chi^2 = \sum_{j=1}^{k}\sum_{i=1}^{r}\frac{(e_{ij}-O_{ij})^2}{e_{ij}} = 20.59$$

المقارنة والقرار:

حيث أن $\chi^2 = 20.59 > (\chi^2_{(0.05,2)} = 7.815$ فإننا نرفض الفرضية الصفرية وبالتالي يوجد دليل على وجود اختلاف في مستوى الرضى عند الفئات الاربعة.

- **اختبار جودة المطابقة Goodness – of fit test**

يستخدم هذا الاختبار في انتماء متغير عشوائي معين يتبع توزيع احتمالي معين كالتوزيع الطبيعي أو توزيع ذي الحدين أو توزيع جاما وهكذا، وهو اختبار ذو اتجاه واحد نحو اليمين وتتلخص فكرة الاختبار في مقارنة عدد المشاهدات O_i التي نسجلها من خلال التجربة مع عدد المشاهدات المتوقع C_i في حالة صحة الفرضية الصفرية.

ويكون المعيار المستخدم في هذا الاختبار هو:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - e_i)^2}{e_i} \sim \chi^2((k-1)-r)$$

حيث r عدد المعالم المجهولة و $e_i = np_i$ وتمثل عدد المشاهدات المتوقع، وتحسب على فرض صحة الفرضية الصفرية، أي باستخدام التوزيع الذي نختبره.

مثال6.13: أجريت دراسة على 320 عائلة لكل منها 5 أطفال وتم تصنيف العائلات من حيث عدد الأطفال الذكور وكانت نتائج الدراسة كما يلي:

X عدد الذكور	0	1	2	3	4	5
O_i عدد العائلات	12	42	92	108	46	20

اختبر الادعاء بأن عدد الأطفال الذكور يتبع توزيع ذي الحدين بمعلمة $p = \frac{1}{2}$ عند مستوى معنوية 0.05

الحل/

يتم حساب القيم المتوقعة باستخدام توزيع ذي الحدين حيث $n = 320$ كالتالي:

$$e_i = n\binom{5}{x}\left(\frac{1}{2}\right)^{x}\left(\frac{1}{2}\right)^{5-x} = n\binom{5}{x}\left(\frac{1}{2}\right)^{5}, x = 0,1,2,3,4,5$$

جدول التوافق

O_i	e_i	$O_i - e_i$	$(O_i - e_i)^2$	$\frac{(O_i - e_i)^2}{e_i}$
12	10	2	4	0.4
42	50	−8	64	1.28
92	100	−8	64	0.64
108	100	8	64	0.64
46	50	−4	16	0.32
20	10	10	100	10
			المجموع	13.28

الفرضيات:

$$H_0{:}X \sim \text{b}(5,\frac{1}{2})$$
$$H_1{:}X\ \ \downarrow\sim \text{b}(5,\frac{1}{2})$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع $\chi^2_{(5)}$ وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاه واحد نحو اليمين إذن تكون القيمة الحرجة:

$$\chi^2_{(0.05,5)} = 11.0705$$

المعيار:

حيث أن التوزيع المطلوب هو التوزيع المنتظم وهو من النوع المنفصل فيكون المعيار هو:

$$\chi^2 = \sum_{i=1}^{n}\frac{(O_i - e_i)^2}{e_i} = 13.28$$

المقارنة والقرار:

وحيث أن $(\chi^2_{(0.05,5)} = 11.0705) > \chi^2 = 13.28$ فإننا نرفض الفرضية الصفرية وبالتالي توزيع الأطفال الذكور لا يتبع التوزيع المذكور.

مثال6.14: يمثل الجدول التالي التوزيع التكراري لعلامات 200 طالباً في مساق التفاضل والتكامل في الكلية.

الفئات G_i	التكرار O_i
< 50	20
$[50 - 59]$	50
$[60 - 69]$	30
$[70 - 79]$	45
$[80 - 89]$	35
≥ 90	20
المجموع	$n = 200$

اختبر الفرضية التي تدعي أن البيانات المعطاة تتبع توزيعاً طبيعياً وسطه 70 وانحرافه المعياري 10 وذلك عند مستوى معنوية 0.05

الحل/

نحول الحدود الفعلية للفئات للدرجة المعيارية باستخدام العلاقة $\boldsymbol{z_i} = \frac{\mathbf{UB}-\boldsymbol{\mu}}{\boldsymbol{\sigma}}$ ثم نحسب احتمال كل فئة كما هو موضح بالجدول التالي:

الحدود الفعلية للفئات UB	z_i	$P(z < z_i)$	$f(z_i)$	الفئة
49.5	−2.05	0.0202	−	< 50
59.5	−1.05	0.1469	0.1267	$[50 - 59]$
69.5	−0.05	0.4801	0.3332	$[60 - 69]$
79.5	0.95	0.8289	0.3488	$[70 - 79]$
89.5	1.95	0.9744	0.1455	$[80 - 89]$
		0.0256	−	≥ 90

حيث احتمال الفئة يحسب من العلاقة $f(z_i) = P(z < z_i) - P(z < z_{i-1})$

جدول التوافق

O_i	$e_i = nf(z_i)$	$O_i - e_i$	$(O_i - e_i)^2$	$\frac{(O_i - e_i)^2}{e_i}$
20	4.04	15.96	254.7216	63.05
50	25.34	24.66	608.1156	24
30	66.64	−36.64	1342.49	20.15
45	69.76	−24.76	613.0576	8.79
35	29.1	5.9	34.81	1.2
			المجموع	117.19

الفرضيات:

$$H_0{:}X \sim \mathrm{N}(70,100)$$
$$H_1{:}X \not\sim \mathrm{N}(70,100)$$

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع $\chi^2_{(5)}$ وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاه واحد نحو اليمين إذن تكون القيمة الحرجة:

$$\chi^2_{(0.05,5)} = 11.070$$

المعيار:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - e_i)^2}{e_i} = 117.19$$

المقارنة والقرار:

وحيث أن $(\chi^2_{(0.05,5)} = 11.070) > \chi^2 = 117.19$ فإننا نرفض الفرضية الصفرية وبالتالي مجتمع الدراسة لا يتبع التوزيع الطبيعي بوسط 70 وانحراف معياري10.

سادساً/ اختبار فريدمان Fred Mann Test

يعتبر هذا الاختبار تعميماً لاختبار ولكيكسون للعينات المرتبطة وهو مشابه لتحليل التباين ويعمل على المقارنة من حيث الفروق بين الطرق، واختبار الفروق بين المتوسطات وهو اختبار ذو اتجاه واحد نحو اليمين.

بفرض أن لدينا k طريقة معالجة هي $T_1, T_2, \dots T_k$ وتم إجراء التجربة ورصد الاستجابات حيث عدد الاستجابات المرصودة لكل معالجة يساوي n، نقوم بترتيب الاستجابات في كل عينة على حده وتحديد رتب الاستجابات ثم نجمع رتب كل عينة على حده حيث نرمز لمجموع رتب العينة j بالرمز R_j ، ويكون المعيار المستخدم هنا هو:

$$\chi^2 = \frac{12}{nk(k+1)} \sum_{j=1}^{k} {R_j}^2 - 3n(k+1) \sim \chi^2(k-1)$$

والمثال التالي يوضح كيفية تطبيق الاختبار.

مثال6.15: يمثل الجدول التالي زمن الشفاء من مرض معين عند تناول المرضى ثلاثة انواع من الأدوية. هل تستطيع أن تستنتج أن هناك فرقاً بين أنواع الأدوية الثلاثة على مستوى دلالة 0.05

	T_1	T_2	T_3
1	10	11	15
2	10	15	20
3	11	15	12
4	8	12	10
5	7	12	9
6	15	10	16
7	14	12	18
8	10	14	17
9	9	9	12
10	10	14	16

الحل/

الجدول التالي يبين رتب الاستجابات في كل عينة ومجموع هذه الرتب.

	T_1	T_2	T_3
1	1	2	3
2	1	2	3
3	1	3	2
4	1	3	2
5	1	3	2
6	2	1	3
7	2	1	3
8	2	2	3
9	1.5	1.5	3
10	1	2	3
المجموع	13.5	20.5	27

الفرضيات:

H_0: لا يوجد فرق بين الأدوية الثلاثة

H_1: يوجد فرق بين الأدوية الثلاثة

التوزيع و القيم الحرجة المعيارية:

التوزيع المستخدم هو توزيع $\chi^2_{(2)}$ وحيث أن $\alpha = 0.05$ والاختبار ذو اتجاه واحد نحو اليمين، إذن تكون القيمة الحرجة $\chi^2_{(0.05,2)} = 5.99$

المعيار:

$$\chi^2 = \frac{12}{nk(k+1)} \sum_{j=1}^{k} {R_j}^2 - 3n(k+1)$$

$$\chi^2 = \frac{12}{10 \times 3(4)} [13.5^2 + 20.5^2 + 27^2] - 3 \times 10(4) = 13.15$$

المقارنة والقرار: وحيث أن $\chi^2 = 11.15 > (\chi^2_{(0.05,2)} = 5.99)$ فإننا نرفض الفرضية الصفرية وبالتالي يوجد فرق بين الأدوية الثلاثة.

تمارين:

1- يحتفظ مدرب سباحة بالأرقام التي يسجلها أحد السباحين في قطع مسافة 50 ياردة. يدعي السباح أن وسيط الزمن الذي يلزمه لقطع المسافة لا يزيد عن 24.8 ثانية إذا كانت الارقام التالية تمثل زمن قطع السباح للمسافة في 17 يوماً:

25.6 24.5 24.4 25.3 26.3 24.2 24.3 24.6 24.5 24.7 26.7
24.7 24.7 24.6 25.7 24.2 24.3

استخدم اختبار الاشارة لاختبار ادعاء السباح على مستوى معنوية 0.1

2- لمناقشة سبل زيادة الارباح في احدى المصانع طرح مجلس الادارة سؤالين على عينة من العاملين عن افضل الطرق، وبعد فرز الاجابات وجد أن 60 عاملاً يفضلون زيادة الدعاية و الاعلانات التسويقية، و أن 40 يفضلون تخفيض أسعار المنتجات لجذب الزبائن. حدد إذا ما كان هناك أفضلية لأحد الاقتراحين على الآخر عند مستوى معنوية 0.1

3- الجدول التالي يبين تقييم أداء عمل بطريقتين، اختبر عند معنوية 0.05 وجود فرق بين الطريقتين باستخدام اختبار اشارة الرتب.

الاولى	167	149	137	178	179	155	164	104	151	151	150
الثانية	98	127	140	103	116	105	100	95	131	132	160

4- الجدول التالي يبين عدد الطلبة الناجحين وعدد الطلبة الراسبين في مساقي الرياضيات والفيزياء و الكيمياء ف الكلية:

	الفيزياء	الكيماء	الرياضيات
الناجحون	65	50	45
الراسبون	15	20	5

اختبر وجود فروق معنوية بين نسب النجاح الثلاثة على مستوى معنوية 0.05

5- تقدم 15 طالباً لامتحانين متشابهين الاول في قاعة غير مكيفة و الثاني في قاعة مكيفة وكانت درجاتهم كما يلي:

الطالب	درجة الاختبار الاول	درجة الاختبار الثاني
1	52	49
2	90	94
3	63	60
4	74	78
5	87	93
6	77	77
7	92	93
8	77	74
9	94	78
10	94	93
11	67	78
12	86	89
13	78	92
14	80	82
15	57	68

اختبر تحسن أداء الطلبة في قاعة مكيفة عند مستوى معنوية 0.1

6- القيت اربع قطع نقدية 1000 مرة وفي كل مرة يتم رصد عدد الصور التي ظهرت وكانت النتيجة في نهاية التجربة كما يلي:

عدد الصور X	0	1	2	3	4
عدد ا لمشاهدات	65	215	400	240	80

اختبر الفرضية $H_0: X \sim b(4, \frac{1}{2})$ عند مستوى معنوية 0.05

7- قامت احدى الشركات بتدريب بعض عمالها على استخدام آلات جديدة، واستخدمت لهذا الغرض اسلوبين للتدريب، ورصد الزمن الذي استغرقه المتدربين لاكتساب المهارة وكان كما يلي:

البرنامج الاول	البرنامج الثاني
40	29
44	27
33	32
26	25
31	27
29	28
34	31
31	23
38	37
33	28
42	22
35	31
	24

اختبر عند مستوى معنوية 0.05 أن البرنامج الثاني أكثر فعالية من البرنامج الاول.

8- اخذت عينة من 500 سائق وتم تصنيفهم حسب العمر وحسب عدد حوادث السير التي اشتركوا فيها في الاعوام الثلاثة الماضية، و كانت النتائج كما يلي:

		العمر بالسنوات		
		دون 30	30–40	اكبر من 40
عدد ا لحوادث	0	61	109	180
	1	27	25	48
	اكثر من 1	12	16	22

اختبر الفرضية المبدئية القائلة "إن عدد الحوادث لا يعتمد على عمر السائق" عند دلالة 0.05

9- يبين الجدول التالي عدد الطلبات على سلعة معينة، وعند رصد الطلبات على سلعة في 200 يوم كانت النتائج كما يلي:

عدد ال طلبات ***X***	0	1	2	3	4	5	6	اكبر من 6
عدد المشاهدات	11	28	43	47	32	28	7	4

اختبر الفرضية $H_0{:}X \sim \mathrm{PO}(3)$ عند مستوى معنوية 0.05

10- اجريت دراسة لمعرفة رأي السيدات ورأي الرجال حول موضوع التوسع في التعليم الاكاديمي النظري و رصدت النتائج في الجدول التالي:

	مع التوسع	ضد التوسع	لا رأي
رجال	86	34	30
سيدات	114	16	20

هل هناك فرق بين رأي السيدات ورأي الرجال حول موضوع التوسع في التعليم الاكاديمي النظري على مستوى دلالة 0.05

11- اختبر الفرضية "لا يوجد فرق بين طرق المعالجة $\mathrm{T_3}, \mathrm{T_2}, \mathrm{T_1}$ على مستوى دلالة 0.05 للبيانات في الجدول التالي:

	$\mathrm{T_1}$	$\mathrm{T_2}$	$\mathrm{T_3}$
1	10	18	7
2	12	19	8
3	15	17	16
4	13	14	12
5	15	20	17
6	12	15	10
7	11	7	6
8	13	18	11
9	15	19	11
10	7	13	12
11	12	13	18
12	10	8	8

12-في دراسة لبيان العلاقة بين دخل الاسرة و انفاقها، سحبت عينة عشوائية بحجم 200 اسرة وكانت النتائج كما يلي:

الدخل / الانفاق	100 -200	200 - 400
متدني	90	45
عالي	15	50

اختبر استقلالية الدخل عن الانفاق عند مستوى معنوية 0.01

الملاحق والمراجع

$\Phi(z) = P(Z \leq z),\ -3.49 \leq z \leq 0\ ,\quad Z \sim N(0,1)$ جدول (1) التوزيع الطبيعي

z	0.0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3829
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

تابع – جدول (1) التوزيع الطبيعي

$\Phi(z) = P(Z \leq z),\ 0 \leq z \leq 3.49\ ,\quad Z \sim N(0, 1)$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

جدول (2) توزيع مربع كاي

$$P(X \geq \mathrm{a}), \qquad X \sim x^2(\alpha, \vartheta)$$

ϑ \ α	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	–	–	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

$P(X \geq a)$, $X \sim t(\vartheta)$ جدول (3) توزيع t

ϑ \ α	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.076	31.821	63.657	318.310	636.620
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

جدول (4) توزيع F

$P(X \geq a), \quad X \sim F(0.05, \vartheta_1, \vartheta_2)$

ϑ_2 \ ϑ_1	1	2	3	4	5	6	7	8	9	10	12	15
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75
inf	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67

تابع - جدول (4) توزيع F

$P(X \geq a), \quad X \sim F(0.05, \vartheta_1, \vartheta_2)$

ϑ_2 \ ϑ_1	20	24	30	40	60	120	inf
1	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.66	1.61	1.55	1.50	1.43	1.35	1.25
inf	1.57	1.52	1.46	1.39	1.32	1.22	1.00

تابع - جدول (4) توزيع F

$P(X \geq a), \quad X \sim F(0.1, \vartheta_1, \vartheta_2)$

ϑ_2 \ ϑ_1	1	2	3	4	5	6	7	8	9	10	12	15
1	39.86	49.5	53.59	55.83	57.24	58.2	58.91	59.44	59.86	60.19	60.71	61.22
2	8.53	9	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.2
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.9	3.87
5	4.06	3.78	3.62	3.52	3.45	3.4	3.37	3.34	3.32	3.3	3.27	3.24
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.9	2.87
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.7	2.67	2.63
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.5	2.46
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.3	2.27	2.25	2.21	2.17
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.1
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.2	2.16	2.14	2.1	2.05
14	3.1	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.1	2.05	2.01
15	3.07	2.7	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94
17	3.03	2.64	2.44	2.31	2.22	2.15	2.1	2.06	2.03	2	1.96	1.91
18	3.01	2.62	2.42	2.29	2.2	2.13	2.08	2.04	2	1.98	1.93	1.89
19	2.99	2.61	2.4	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2	1.96	1.94	1.89	1.84
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.9	1.86	1.81
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.8
24	2.93	2.54	2.33	2.19	2.1	2.04	1.98	1.94	1.91	1.88	1.83	1.78
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76
27	2.9	2.51	2.3	2.17	2.07	2	1.95	1.91	1.87	1.85	1.8	1.75
28	2.89	2.5	2.29	2.16	2.06	2	1.94	1.9	1.87	1.84	1.79	1.74
29	2.89	2.5	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72
40	2.84	2.44	2.23	2.09	2	1.93	1.87	1.83	1.79	1.76	1.71	1.66
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.6
120	2.75	2.35	2.13	1.99	1.9	1.82	1.77	1.72	1.68	1.65	1.6	1.55
inf	2.71	2.3	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.6	1.55	1.49

تابع - جدول (4) توزيع F

$$P(X \geq a), \qquad X \sim F(0.1, \vartheta_1, \vartheta_2)$$

ϑ_2 \ ϑ_1	20	24	30	40	60	120	inf
1	61.74	62	62.26	62.53	62.79	63.06	63.33
2	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	3.84	3.83	3.82	3.8	3.79	3.78	3.76
5	3.21	3.19	3.17	3.16	3.14	3.12	3.11
6	2.84	2.82	2.8	2.78	2.76	2.74	2.72
7	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	2.42	2.4	2.38	2.36	2.34	2.32	2.29
9	2.3	2.28	2.25	2.23	2.21	2.18	2.16
10	2.2	2.18	2.16	2.13	2.11	2.08	2.06
11	2.12	2.1	2.08	2.05	2.03	2	1.97
12	2.06	2.04	2.01	1.99	1.96	1.93	1.9
13	2.01	1.98	1.96	1.93	1.9	1.88	1.85
14	1.96	1.94	1.91	1.89	1.86	1.83	1.8
15	1.92	1.9	1.87	1.85	1.82	1.79	1.76
16	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	1.84	1.81	1.78	1.75	1.72	1.69	1.66
19	1.81	1.79	1.76	1.73	1.7	1.67	1.63
20	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	1.76	1.73	1.7	1.67	1.64	1.6	1.57
23	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	1.73	1.7	1.67	1.64	1.61	1.57	1.53
25	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	1.71	1.68	1.65	1.61	1.58	1.54	1.5
27	1.7	1.67	1.64	1.6	1.57	1.53	1.49
28	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	1.67	1.64	1.61	1.57	1.54	1.5	1.46
40	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	1.54	1.51	1.48	1.44	1.4	1.35	1.29
120	1.48	1.45	1.41	1.37	1.32	1.26	1.19
inf	1.42	1.38	1.34	1.3	1.24	1.17	1

تابع - جدول (4) توزيع F

$P(\mathrm{X} \geq \mathrm{a}), \qquad \mathrm{X}\sim\mathrm{F}(0.025, \vartheta_1, \vartheta_2)$

ϑ_2 \ ϑ_1	1	2	3	4	5	6	7	8	9	10	12	15
1	647.79	799.5	864.16	899.58	921.85	937.11	948.22	956.66	963.28	968.63	976.71	984.87
2	38.51	39	39.17	39.25	39.3	39.33	39.36	39.37	39.39	39.4	39.41	39.43
3	17.44	16.04	15.44	15.1	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25
4	12.22	10.65	9.98	9.6	9.36	9.2	9.07	8.98	8.9	8.84	8.75	8.66
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43
6	8.81	7.26	6.6	6.23	5.99	5.82	5.7	5.6	5.52	5.46	5.37	5.27
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.9	4.82	4.76	4.67	4.57
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.3	4.2	4.1
9	7.21	5.71	5.08	4.72	4.48	4.32	4.2	4.1	4.03	3.96	3.87	3.77
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33
12	6.55	5.1	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18
13	6.41	4.97	4.35	4	3.77	3.6	3.48	3.39	3.31	3.25	3.15	3.05
14	6.3	4.86	4.24	3.89	3.66	3.5	3.38	3.29	3.21	3.15	3.05	2.95
15	6.2	4.77	4.15	3.8	3.58	3.41	3.29	3.2	3.12	3.06	2.96	2.86
16	6.12	4.69	4.08	3.73	3.5	3.34	3.22	3.12	3.05	2.99	2.89	2.79
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72
18	5.98	4.56	3.95	3.61	3.38	3.22	3.1	3.01	2.93	2.87	2.77	2.67
19	5.92	4.51	3.9	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.8	2.73	2.64	2.53
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.7	2.6	2.5
23	5.75	4.35	3.75	3.41	3.18	3.02	2.9	2.81	2.73	2.67	2.57	2.47
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.7	2.64	2.54	2.44
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41
26	5.66	4.27	3.67	3.33	3.1	2.94	2.82	2.73	2.65	2.59	2.49	2.39
27	5.63	4.24	3.65	3.31	3.08	2.92	2.8	2.71	2.63	2.57	2.47	2.36
28	5.61	4.22	3.63	3.29	3.06	2.9	2.78	2.69	2.61	2.55	2.45	2.34
29	5.59	4.2	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31
40	5.42	4.05	3.46	3.13	2.9	2.74	2.62	2.53	2.45	2.39	2.29	2.18
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06
120	5.15	3.8	3.23	2.89	2.67	2.52	2.39	2.3	2.22	2.16	2.05	1.95
inf	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83

تابع - جدول (4) توزيع F

$$P(X \geq a), \quad X \sim F(0.025, \vartheta_1, \vartheta_2)$$

ϑ_2 \ ϑ_1	20	24	30	40	60	120	inf
1	993.1	997.25	1001.41	1005.6	1009.8	1014.02	1018.26
2	39.45	39.46	39.47	39.47	39.48	39.49	39.5
3	14.17	14.12	14.08	14.04	13.99	13.95	13.9
4	8.56	8.51	8.46	8.41	8.36	8.31	8.26
5	6.33	6.28	6.23	6.18	6.12	6.07	6.02
6	5.17	5.12	5.07	5.01	4.96	4.9	4.85
7	4.47	4.42	4.36	4.31	4.25	4.2	4.14
8	4	3.95	3.89	3.84	3.78	3.73	3.67
9	3.67	3.61	3.56	3.51	3.45	3.39	3.33
10	3.42	3.37	3.31	3.26	3.2	3.14	3.08
11	3.23	3.17	3.12	3.06	3	2.94	2.88
12	3.07	3.02	2.96	2.91	2.85	2.79	2.73
13	2.95	2.89	2.84	2.78	2.72	2.66	2.6
14	2.84	2.79	2.73	2.67	2.61	2.55	2.49
15	2.76	2.7	2.64	2.59	2.52	2.46	2.4
16	2.68	2.63	2.57	2.51	2.45	2.38	2.32
17	2.62	2.56	2.5	2.44	2.38	2.32	2.25
18	2.56	2.5	2.45	2.38	2.32	2.26	2.19
19	2.51	2.45	2.39	2.33	2.27	2.2	2.13
20	2.46	2.41	2.35	2.29	2.22	2.16	2.09
21	2.42	2.37	2.31	2.25	2.18	2.11	2.04
22	2.39	2.33	2.27	2.21	2.15	2.08	2
23	2.36	2.3	2.24	2.18	2.11	2.04	1.97
24	2.33	2.27	2.21	2.15	2.08	2.01	1.94
25	2.3	2.24	2.18	2.12	2.05	1.98	1.91
26	2.28	2.22	2.16	2.09	2.03	1.95	1.88
27	2.25	2.19	2.13	2.07	2	1.93	1.85
28	2.23	2.17	2.11	2.05	1.98	1.91	1.83
29	2.21	2.15	2.09	2.03	1.96	1.89	1.81
30	2.2	2.14	2.07	2.01	1.94	1.87	1.79
40	2.07	2.01	1.94	1.88	1.8	1.72	1.64
60	1.94	1.88	1.82	1.74	1.67	1.58	1.48
120	1.82	1.76	1.69	1.61	1.53	1.43	1.31
inf	1.71	1.64	1.57	1.48	1.39	1.27	1

تابع - جدول (4) توزيع F

$P(X \geq a), \quad X \sim F(0.01, \vartheta_1, \vartheta_2)$

ϑ_2 \ ϑ_1	1	2	3	4	5	6	7	8	9	10	12	15
1	4052.18	4999.5	5403.35	5624.58	5763.65	5858.99	5928.36	5981.07	6022.47	6055.85	6106.32	6157.29
2	98.5	99	99.17	99.25	99.3	99.33	99.36	99.37	99.39	99.4	99.42	99.43
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87
4	21.2	18	16.69	15.98	15.52	15.21	14.98	14.8	14.66	14.55	14.37	14.2
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72
6	13.75	10.93	9.78	9.15	8.75	8.47	8.26	8.1	7.98	7.87	7.72	7.56
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52
9	10.56	8.02	6.99	6.42	6.06	5.8	5.61	5.47	5.35	5.26	5.11	4.96
10	10.04	7.56	6.55	5.99	5.64	5.39	5.2	5.06	4.94	4.85	4.71	4.56
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.4	4.25
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.5	4.39	4.3	4.16	4.01
13	9.07	6.7	5.74	5.21	4.86	4.62	4.44	4.3	4.19	4.1	3.96	3.82
14	8.86	6.52	5.56	5.04	4.7	4.46	4.28	4.14	4.03	3.94	3.8	3.66
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4	3.9	3.81	3.67	3.52
16	8.53	6.23	5.29	4.77	4.44	4.2	4.03	3.89	3.78	3.69	3.55	3.41
17	8.4	6.11	5.19	4.67	4.34	4.1	3.93	3.79	3.68	3.59	3.46	3.31
18	8.29	6.01	5.09	4.58	4.25	4.02	3.84	3.71	3.6	3.51	3.37	3.23
19	8.19	5.93	5.01	4.5	4.17	3.94	3.77	3.63	3.52	3.43	3.3	3.15
20	8.1	5.85	4.94	4.43	4.1	3.87	3.7	3.56	3.46	3.37	3.23	3.09
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.4	3.31	3.17	3.03
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98
23	7.88	5.66	4.77	4.26	3.94	3.71	3.54	3.41	3.3	3.21	3.07	2.93
24	7.82	5.61	4.72	4.22	3.9	3.67	3.5	3.36	3.26	3.17	3.03	2.89
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	2.99	2.85
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.82
27	7.68	5.49	4.6	4.11	3.79	3.56	3.39	3.26	3.15	3.06	2.93	2.78
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.9	2.75
29	7.6	5.42	4.54	4.05	3.73	3.5	3.33	3.2	3.09	3.01	2.87	2.73
30	7.56	5.39	4.51	4.02	3.7	3.47	3.3	3.17	3.07	2.98	2.84	2.7
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.8	2.67	2.52
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.5	2.35
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19
inf	6.64	4.61	3.78	3.32	3.02	2.8	2.64	2.51	2.41	2.32	2.19	2.04

تابع - جدول (4) توزيع F

$P(X \geq a), \quad X \sim F(0.01, \vartheta_1, \vartheta_2)$

ϑ_2 \ ϑ_1	20	24	30	40	60	120	inf
1	6208.73	6234.63	6260.65	6286.78	6313.03	6339.39	6365.86
2	99.45	99.46	99.47	99.47	99.48	99.49	99.5
3	26.69	26.6	26.51	26.41	26.32	26.22	26.13
4	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	9.55	9.47	9.38	9.29	9.2	9.11	9.02
6	7.4	7.31	7.23	7.14	7.06	6.97	6.88
7	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	5.36	5.28	5.2	5.12	5.03	4.95	4.86
9	4.81	4.73	4.65	4.57	4.48	4.4	4.31
10	4.41	4.33	4.25	4.17	4.08	4	3.91
11	4.1	4.02	3.94	3.86	3.78	3.69	3.6
12	3.86	3.78	3.7	3.62	3.54	3.45	3.36
13	3.67	3.59	3.51	3.43	3.34	3.26	3.17
14	3.51	3.43	3.35	3.27	3.18	3.09	3
15	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	3.26	3.18	3.1	3.02	2.93	2.85	2.75
17	3.16	3.08	3	2.92	2.84	2.75	2.65
18	3.08	3	2.92	2.84	2.75	2.66	2.57
19	3	2.93	2.84	2.76	2.67	2.58	2.49
20	2.94	2.86	2.78	2.7	2.61	2.52	2.42
21	2.88	2.8	2.72	2.64	2.55	2.46	2.36
22	2.83	2.75	2.67	2.58	2.5	2.4	2.31
23	2.78	2.7	2.62	2.54	2.45	2.35	2.26
24	2.74	2.66	2.58	2.49	2.4	2.31	2.21
25	2.7	2.62	2.54	2.45	2.36	2.27	2.17
26	2.66	2.59	2.5	2.42	2.33	2.23	2.13
27	2.63	2.55	2.47	2.38	2.29	2.2	2.1
28	2.6	2.52	2.44	2.35	2.26	2.17	2.06
29	2.57	2.5	2.41	2.33	2.23	2.14	2.03
30	2.55	2.47	2.39	2.3	2.21	2.11	2.01
40	2.37	2.29	2.2	2.11	2.02	1.92	1.81
60	2.2	2.12	2.03	1.94	1.84	1.73	1.6
120	2.04	1.95	1.86	1.76	1.66	1.53	1.38
inf	1.88	1.79	1.7	1.59	1.47	1.33	1

جدول (5)   دونيت

n	α	Number of Groups, Including Control Group								
		2	3	4	5	6	7	8	9	10
5	0.05	2.57	3.03	3.29	3.48	3.62	3.73	3.82	3.90	3.97
	0.01	4.03	4.63	4.98	5.22	5.41	5.56	5.69	5.80	5.89
6	0.05	2.45	2.86	3.10	3.26	3.39	3.49	3.57	3.64	3.71
	0.01	3.71	4.21	4.51	4.71	4.87	5.00	5.10	5.20	5.28
7	0.05	2.36	2.75	2.97	3.12	3.24	3.33	3.41	3.47	3.53
	0.01	3.50	3.95	4.21	4.39	4.53	4.64	4.74	4.82	4.89
8	0.05	2.31	2.67	2.88	3.02	3.13	3.22	3.29	3.35	3.41
	0.01	3.36	3.77	4.00	4.17	4.29	4.40	4.48	4.56	4.62
9	0.05	2.26	2.61	2.81	2.95	3.05	3.14	3.20	3.26	3.32
	0.01	3.25	3.63	3.85	4.01	4.12	4.22	4.30	4.37	4.43
10	0.05	2.23	2.57	2.76	2.89	2.99	3.07	3.14	3.19	3.24
	0.01	3.17	3.53	3.74	3.88	3.99	4.08	4.16	4.22	4.28
11	0.05	2.20	2.53	2.72	2.84	2.94	3.02	3.08	3.14	3.19
	0.01	3.11	3.45	3.65	3.79	3.89	3.98	4.05	4.11	4.16
12	0.05	2.18	2.50	2.68	2.81	2.90	2.98	3.04	3.09	3.14
	0.01	3.05	3.39	3.58	3.71	3.81	3.89	3.96	4.02	4.07
13	0.05	2.16	2.48	2.65	2.78	2.87	2.94	3.00	3.06	3.10
	0.01	3.01	3.33	3.52	3.65	3.74	3.82	3.89	3.94	3.99
14	0.05	2.14	2.46	2.63	2.75	2.84	2.91	2.97	3.02	3.07
	0.01	2.98	3.29	3.47	3.59	3.69	3.76	3.83	3.88	3.93
15	0.05	2.13	2.44	2.61	2.73	2.82	2.89	2.95	3.00	3.04
	0.01	2.95	3.25	3.43	3.55	3.64	3.71	3.78	3.83	3.88
16	0.05	2.12	2.42	2.59	2.71	2.80	2.87	2.92	2.97	3.02
	0.01	2.92	3.22	3.39	3.51	3.60	3.67	3.73	3.78	3.83
17	0.05	2.11	2.41	2.58	2.69	2.78	2.85	2.90	2.95	3.00
	0.01	2.90	3.19	3.36	3.47	3.56	3.63	3.69	3.74	3.79
18	0.05	2.10	2.40	2.56	2.68	2.76	2.83	2.89	2.94	2.98
	0.01	2.88	3.17	3.33	3.44	3.53	3.60	3.66	3.71	3.75
19	0.05	2.09	2.39	2.55	2.66	2.75	2.81	2.87	2.92	2.96
	0.01	2.86	3.15	3.31	3.42	3.50	3.57	3.63	3.68	3.72
20	0.05	2.09	2.38	2.54	2.65	2.73	2.80	2.86	2.90	2.95
	0.01	2.85	3.13	3.29	3.40	3.48	3.55	3.60	3.65	3.69
24	0.05	2.06	2.35	2.51	2.61	2.70	2.76	2.81	2.86	2.90
	0.01	2.80	3.07	3.22	3.32	3.40	3.47	3.52	3.57	3.61
30	0.05	2.04	2.32	2.47	2.58	2.66	2.72	2.77	2.82	2.86
	0.01	2.75	3.01	3.15	3.25	3.33	3.39	3.44	3.49	3.52
40	0.05	2.02	2.29	2.44	2.54	2.62	2.68	2.73	2.77	2.81
	0.01	2.70	2.95	3.09	3.19	3.26	3.32	3.37	3.41	3.44
60	0.05	2.00	2.27	2.41	2.51	2.58	2.64	2.69	2.73	2.77
	0.01	2.66	2.90	3.03	3.12	3.19	3.25	3.29	3.33	3.37

جدول (6) توزيع ذي الحدين

Cumulative Binomial Distribution Table

n=1

x	p = 0.10	p = 0.20	p = 0.30	p = 0.40	p = 0.50	p = 0.60	p = 0.70	p = 0.80	p = 0.90
0	0.9000	0.8000	0.7000	0.6000	0.5000	0.4000	0.3000	0.2000	0.1000
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

n=2

x	p = 0.10	p = 0.20	p = 0.30	p = 0.40	p = 0.50	p = 0.60	p = 0.70	p = 0.80	p = 0.90
0	0.9801	0.9409	0.9025	0.8649	0.8281	0.1600	0.0900	0.0400	0.0100
1	0.9999	0.9991	0.9975	0.9951	0.9919	0.6400	0.5100	0.3600	0.1900
2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

n=3

x	p = 0.10	p = 0.20	p = 0.30	p = 0.40	p = 0.50	p = 0.60	p = 0.70	p = 0.80	p = 0.90
0	0.7290	0.5120	0.3430	0.2160	0.1250	0.0640	0.0270	0.0080	0.0010
1	0.9720	0.8960	0.7840	0.6480	0.5000	0.3520	0.2160	0.1040	0.0280
2	0.9990	0.9920	0.9730	0.9360	0.8750	0.7840	0.6570	0.4880	0.2710
3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

n=4

x	p = 0.10	p = 0.20	p = 0.30	p = 0.40	p = 0.50	p = 0.60	p = 0.70	p = 0.80	p = 0.90
0	0.6561	0.4096	0.2401	0.1296	0.0625	0.0256	0.0081	0.0016	0.0001
1	0.9477	0.8192	0.6517	0.4752	0.3125	0.1792	0.0837	0.0272	0.0037
2	0.9963	0.9728	0.9163	0.8208	0.6875	0.5248	0.3483	0.1808	0.0523
3	0.9999	0.9984	0.9919	0.9744	0.9375	0.8704	0.7599	0.5904	0.3439
4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

n=5

x	p = 0.10	p = 0.20	p = 0.30	p = 0.40	p = 0.50	p = 0.60	p = 0.70	p = 0.80	p = 0.90
0	0.5905	0.3277	0.1681	0.0778	0.0313	0.0102	0.0024	0.0003	0.0000
1	0.9185	0.7373	0.5282	0.3370	0.1875	0.0870	0.0308	0.0067	0.0005
2	0.9914	0.9421	0.8369	0.6826	0.5000	0.3174	0.1631	0.0579	0.0086
3	0.9995	0.9933	0.9692	0.9130	0.8125	0.6630	0.4718	0.2627	0.0815
4	1.0000	0.9997	0.9976	0.9898	0.9688	0.9222	0.8319	0.6723	0.4095
5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

n=6

x	p = 0.10	p = 0.20	p = 0.30	p = 0.40	p = 0.50	p = 0.60	p = 0.70	p = 0.80	p = 0.90
0	0.5314	0.2621	0.1176	0.0467	0.0156	0.0041	0.0007	0.0001	0.0000
1	0.8857	0.6554	0.4202	0.2333	0.1094	0.0410	0.0109	0.0016	0.0001
2	0.9842	0.9011	0.7443	0.5443	0.3438	0.1792	0.0705	0.0170	0.0013
3	0.9987	0.9830	0.9295	0.8208	0.6563	0.4557	0.2557	0.0989	0.0159
4	0.9999	0.9984	0.9891	0.9590	0.8906	0.7667	0.5798	0.3446	0.1143
5	1.0000	0.9999	0.9993	0.9959	0.9844	0.9533	0.8824	0.7379	0.4686
6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

تابع - جدول (6) توزيع ذي الحدين

	n=7								
x	**p = 0.10**	**p = 0.20**	**p = 0.30**	**p = 0.40**	**p = 0.50**	**p = 0.60**	**p = 0.70**	**p = 0.80**	**p = 0.90**
0	0.4783	0.2097	0.0824	0.0280	0.0078	0.0016	0.0002	0.0000	0.0000
1	0.8503	0.5767	0.3294	0.1586	0.0625	0.0188	0.0038	0.0004	0.0000
2	0.9743	0.8520	0.6471	0.4199	0.2266	0.0963	0.0288	0.0047	0.0002
3	0.9973	0.9667	0.8740	0.7102	0.5000	0.2898	0.1260	0.0333	0.0027
4	0.9998	0.9953	0.9712	0.9037	0.7734	0.5801	0.3529	0.1480	0.0257
5	1.0000	0.9996	0.9962	0.9812	0.9375	0.8414	0.6706	0.4233	0.1497
6	1.0000	1.0000	0.9998	0.9984	0.9922	0.9720	0.9176	0.7903	0.5217
7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	n=8								
x	**p = 0.10**	**p = 0.20**	**p = 0.30**	**p = 0.40**	**p = 0.50**	**p = 0.60**	**p = 0.70**	**p = 0.80**	**p = 0.90**
0	0.4305	0.1678	0.0576	0.0168	0.0039	0.0007	0.0001	0.0000	0.0000
1	0.8131	0.5033	0.2553	0.1064	0.0352	0.0085	0.0013	0.0001	0.0000
2	0.9619	0.7969	0.5518	0.3154	0.1445	0.0498	0.0113	0.0012	0.0000
3	0.9950	0.9437	0.8059	0.5941	0.3633	0.1737	0.0580	0.0104	0.0004
4	0.9996	0.9896	0.9420	0.8263	0.6367	0.4059	0.1941	0.0563	0.0050
5	1.0000	0.9988	0.9887	0.9502	0.8555	0.6846	0.4482	0.2031	0.0381
6	1.0000	0.9999	0.9987	0.9915	0.9648	0.8936	0.7447	0.4967	0.1869
7	1.0000	1.0000	0.9999	0.9993	0.9961	0.9832	0.9424	0.8322	0.5695
8	1.0000	1.0000	1.0000	1.0000		1.0000	1.0000	1.0000	1.0000

	n=9								
x	**p = 0.10**	**p = 0.20**	**p = 0.30**	**p = 0.40**	**p = 0.50**	**p = 0.60**	**p = 0.70**	**p = 0.80**	**p = 0.90**
0	0.3874	0.1342	0.0404	0.0101	0.0020	0.0003	0.0000	0.0000	0.0000
1	0.7748	0.4362	0.1960	0.0705	0.0195	0.0038	0.0004	0.0000	0.0000
2	0.9470	0.7382	0.4628	0.2318	0.0898	0.0250	0.0043	0.0003	0.0000
3	0.9917	0.9144	0.7297	0.4826	0.2539	0.0994	0.0253	0.0031	0.0001
4	0.9991	0.9804	0.9012	0.7334	0.5000	0.2666	0.0988	0.0196	0.0009
5	0.9999	0.9969	0.9747	0.9006	0.7461	0.5174	0.2703	0.0856	0.0083
6	1.0000	0.9997	0.9957	0.9750	0.9102	0.7682	0.5372	0.2618	0.0530
7	1.0000	1.0000	0.9996	0.9962	0.9805	0.9295	0.8040	0.5638	0.2252
8	1.0000	1.0000	1.0000	0.9997	0.9980	0.9899	0.9596	0.8658	0.6126
9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

تابع - جدول (6) توزيع ذي الحدين

	n=10								
x	**p = 0.10**	**p = 0.20**	**p = 0.30**	**p = 0.40**	**p = 0.50**	**p = 0.60**	**p = 0.70**	**p = 0.80**	**p = 0.90**
0	0.3487	0.1074	0.0282	0.0060	0.0010	0.0001	0.0000	0.0000	0.0000
1	0.7361	0.3758	0.1493	0.0464	0.0107	0.0017	0.0001	0.0000	0.0000
2	0.9298	0.6778	0.3828	0.1673	0.0547	0.0123	0.0016	0.0001	0.0000
3	0.9872	0.8791	0.6496	0.3823	0.1719	0.0548	0.0106	0.0009	0.0000
4	0.9984	0.9672	0.8497	0.6331	0.3770	0.1662	0.0473	0.0064	0.0001
5	0.9999	0.9936	0.9527	0.8338	0.6230	0.3669	0.1503	0.0328	0.0016
6	1.0000	0.9991	0.9894	0.9452	0.8281	0.6177	0.3504	0.1209	0.0128
7	1.0000	0.9999	0.9984	0.9877	0.9453	0.8327	0.6172	0.3222	0.0702
8	1.0000	1.0000	0.9999	0.9983	0.9893	0.9536	0.8507	0.6242	0.2639
9	1.0000	1.0000	1.0000	0.9999	0.9990	0.9940	0.9718	0.8926	0.6513
10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

	n=11								
x	**p = 0.10**	**p = 0.20**	**p = 0.30**	**p = 0.40**	**p = 0.50**	**p = 0.60**	**p = 0.70**	**p = 0.80**	**p = 0.90**
0	0.3138	0.0859	0.0198	0.0036	0.0005	0.0000	0.0000	0.0000	0.0000
1	0.6974	0.3221	0.1130	0.0302	0.0059	0.0007	0.0000	0.0000	0.0000
2	0.9104	0.6174	0.3127	0.1189	0.0327	0.0059	0.0006	0.0000	0.0000
3	0.9815	0.8389	0.5696	0.2963	0.1133	0.0293	0.0043	0.0002	0.0000
4	0.9972	0.9496	0.7897	0.5328	0.2744	0.0994	0.0216	0.0020	0.0000
5	0.9997	0.9883	0.9218	0.7535	0.5000	0.2465	0.0782	0.0117	0.0003
6	1.0000	0.9980	0.9784	0.9006	0.7256	0.4672	0.2103	0.0504	0.0028
7	1.0000	0.9998	0.9957	0.9707	0.8867	0.7037	0.4304	0.1611	0.0185
8	1.0000	1.0000	0.9994	0.9941	0.9673	0.8811	0.6873	0.3826	0.0896
9	1.0000	1.0000	1.0000	0.9993	0.9941	0.9698	0.8870	0.6779	0.3026
10	1.0000	1.0000	1.0000	1.0000	0.9995	0.9964	0.9802	0.9141	0.6862
11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

تابع - جدول (6) توزيع ذي الحدين

	n=12								
x	**p = 0.10**	**p = 0.20**	**p = 0.30**	**p = 0.40**	**p = 0.50**	**p = 0.60**	**p = 0.70**	**p = 0.80**	**p = 0.90**
0	0.2824	0.0687	0.0138	0.0022	0.0002	0.0000	0.0000	0.0000	0.0000
1	0.6590	0.2749	0.0850	0.0196	0.0032	0.0003	0.0000	0.0000	0.0000
2	0.8891	0.5583	0.2528	0.0834	0.0193	0.0028	0.0002	0.0000	0.0000
3	0.9744	0.7946	0.4925	0.2253	0.0730	0.0153	0.0017	0.0001	0.0000
4	0.9957	0.9274	0.7237	0.4382	0.1938	0.0573	0.0095	0.0006	0.0000
5	0.9995	0.9806	0.8822	0.6652	0.3872	0.1582	0.0386	0.0039	0.0001
6	0.9999	0.9961	0.9614	0.8418	0.6128	0.3348	0.1178	0.0194	0.0005
7	1.0000	0.9994	0.9905	0.9427	0.8062	0.5618	0.2763	0.0726	0.0043
8	1.0000	0.9999	0.9983	0.9847	0.9270	0.7747	0.5075	0.2054	0.0256
9	1.0000	1.0000	0.9998	0.9972	0.9807	0.9166	0.7472	0.4417	0.1109
10	1.0000	1.0000	1.0000	0.9997	0.9968	0.9804	0.9150	0.7251	0.3410
11	1.0000	1.0000	1.0000	1.0000	0.9998	0.9978	0.9862	0.9313	0.7176
12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	n=13								
x	**p = 0.10**	**p = 0.20**	**p = 0.30**	**p = 0.40**	**p = 0.50**	**p = 0.60**	**p = 0.70**	**p = 0.80**	**p = 0.90**
0	0.2542	0.0550	0.0097	0.0013	0.0001	0.0000	0.0000	0.0000	0.0000
1	0.6213	0.2336	0.0637	0.0126	0.0017	0.0001	0.0000	0.0000	0.0000
2	0.8661	0.5017	0.2025	0.0579	0.0112	0.0013	0.0001	0.0000	0.0000
3	0.9658	0.7473	0.4206	0.1686	0.0461	0.0078	0.0007	0.0000	0.0000
4	0.9935	0.9009	0.6543	0.3530	0.1334	0.0321	0.0040	0.0002	0.0000
5	0.9991	0.9700	0.8346	0.5744	0.2905	0.0977	0.0182	0.0012	0.0000
6	0.9999	0.9930	0.9376	0.7712	0.5000	0.2288	0.0624	0.0070	0.0001
7	1.0000	0.9988	0.9818	0.9023	0.7095	0.4256	0.1654	0.0300	0.0009
8	1.0000	0.9998	0.9960	0.9679	0.8666	0.6470	0.3457	0.0991	0.0065
9	1.0000	1.0000	0.9993	0.9922	0.9539	0.8314	0.5794	0.2527	0.0342
10	1.0000	1.0000	0.9999	0.9987	0.9888	0.9421	0.7975	0.4983	0.1339
11	1.0000	1.0000	1.0000	0.9999	0.9983	0.9874	0.9363	0.7664	0.3787
12	1.0000	1.0000	1.0000	1.0000	0.9999	0.9987	0.9903	0.9450	0.7458
13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

تابع - جدول (6) توزيع ذي الحدين

	n=14								
x	**p = 0.10**	**p = 0.20**	**p = 0.30**	**p = 0.40**	**p = 0.50**	**p = 0.60**	**p = 0.70**	**p = 0.80**	**p = 0.90**
0	0.2288	0.0440	0.0068	0.0008	0.0001	0.0000	0.0000	0.0000	0.0000
1	0.5846	0.1979	0.0475	0.0081	0.0009	0.0001	0.0000	0.0000	0.0000
2	0.8416	0.4481	0.1608	0.0398	0.0065	0.0006	0.0000	0.0000	0.0000
3	0.9559	0.6982	0.3552	0.1243	0.0287	0.0039	0.0002	0.0000	0.0000
4	0.9908	0.8702	0.5842	0.2793	0.0898	0.0175	0.0017	0.0000	0.0000
5	0.9985	0.9561	0.7805	0.4859	0.2120	0.0583	0.0083	0.0004	0.0000
6	0.9998	0.9884	0.9067	0.6925	0.3953	0.1501	0.0315	0.0024	0.0000
7	1.0000	0.9976	0.9685	0.8499	0.6047	0.3075	0.0933	0.0116	0.0002
8	1.0000	0.9996	0.9917	0.9417	0.7880	0.5141	0.2195	0.0439	0.0015
9	1.0000	1.0000	0.9983	0.9825	0.9102	0.7207	0.4158	0.1298	0.0092
10	1.0000	1.0000	0.9998	0.9961	0.9713	0.8757	0.6448	0.3018	0.0441
11	1.0000	1.0000	1.0000	0.9994	0.9935	0.9602	0.8392	0.5519	0.1584
12	1.0000	1.0000	1.0000	0.9999	0.9991	0.9919	0.9525	0.8021	0.4154
13	1.0000	1.0000	1.0000	1.0000	0.9999	0.9992	0.9932	0.9560	0.7712
14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

	n=15								
x	**p = 0.10**	**p = 0.20**	**p = 0.30**	**p = 0.40**	**p = 0.50**	**p = 0.60**	**p = 0.70**	**p = 0.80**	**p = 0.90**
0	0.2059	0.0352	0.0047	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000
1	0.5490	0.1671	0.0353	0.0052	0.0005	0.0000	0.0000	0.0000	0.0000
2	0.8159	0.3980	0.1268	0.0271	0.0037	0.0003	0.0000	0.0000	0.0000
3	0.9444	0.6482	0.2969	0.0905	0.0176	0.0019	0.0001	0.0000	0.0000
4	0.9873	0.8358	0.5155	0.2173	0.0592	0.0093	0.0007	0.0000	0.0000
5	0.9978	0.9389	0.7216	0.4032	0.1509	0.0338	0.0037	0.0001	0.0000
6	0.9997	0.9819	0.8689	0.6098	0.3036	0.0950	0.0152	0.0008	0.0000
7	1.0000	0.9958	0.9500	0.7869	0.5000	0.2131	0.0500	0.0042	0.0000
8	1.0000	0.9992	0.9848	0.9050	0.6964	0.3902	0.1311	0.0181	0.0003
9	1.0000	0.9999	0.9963	0.9662	0.8491	0.5968	0.2784	0.0611	0.0022
10	1.0000	1.0000	0.9993	0.9907	0.9408	0.7827	0.4845	0.1642	0.0127
11	1.0000	1.0000	0.9999	0.9981	0.9824	0.9095	0.7031	0.3518	0.0556
12	1.0000	1.0000	1.0000	0.9997	0.9963	0.9729	0.8732	0.6020	0.1841
13	1.0000	1.0000	1.0000	1.0000	0.9995	0.9948	0.9647	0.8329	0.4510
14	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9953	0.9648	0.7941
15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

تابع - جدول (6) توزيع ذي الحدين

	n=20								
x	**p = 0.10**	**p = 0.20**	**p = 0.30**	**p = 0.40**	**p = 0.50**	**p = 0.60**	**p = 0.70**	**p = 0.80**	**p = 0.90**
0	0.1216	0.0115	0.0008	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1	0.3917	0.0692	0.0076	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.6769	0.2061	0.0355	0.0036	0.0002	0.0000	0.0000	0.0000	0.0000
3	0.8670	0.4114	0.1071	0.0160	0.0013	0.0000	0.0000	0.0000	0.0000
4	0.9568	0.6296	0.2375	0.0510	0.0059	0.0003	0.0000	0.0000	0.0000
5	0.9887	0.8042	0.4164	0.1256	0.0207	0.0016	0.0000	0.0000	0.0000
6	0.9976	0.9133	0.6080	0.2500	0.0577	0.0065	0.0003	0.0000	0.0000
7	0.9996	0.9679	0.7723	0.4159	0.1316	0.0210	0.0013	0.0000	0.0000
8	0.9999	0.9900	0.8867	0.5956	0.2517	0.0565	0.0051	0.0001	0.0000
9	1.0000	0.9974	0.9520	0.7553	0.4119	0.1275	0.0171	0.0006	0.0000
10	1.0000	0.9994	0.9829	0.8725	0.5881	0.2447	0.0480	0.0026	0.0000
11	1.0000	0.9999	0.9949	0.9435	0.7483	0.4044	0.1133	0.0100	0.0001
12	1.0000	1.0000	0.9987	0.9790	0.8684	0.5841	0.2277	0.0321	0.0004
13	1.0000	1.0000	0.9997	0.9935	0.9423	0.7500	0.3920	0.0867	0.0024
14	1.0000	1.0000	1.0000	0.9984	0.9793	0.8744	0.5836	0.1958	0.0113
15	1.0000	1.0000	1.0000	0.9997	0.9941	0.9490	0.7625	0.3704	0.0432
16	1.0000	1.0000	1.0000	1.0000	0.9987	0.9840	0.8929	0.5886	0.1330
17	1.0000	1.0000	1.0000	1.0000	0.9998	0.9964	0.9645	0.7939	0.3231
18	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9924	0.9308	0.6083
19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9992	0.9885	0.8784
20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

جدول (7) ولكيكسون

$$P(\mathrm{T}^{+} \geq x) = p_r(\mathrm{T}^{+} \leq x^{*})$$

$n=3$			$n=4$			$n=5$			$n=6$		
x	P	x^*	x	P	x^*	x	P	x^*	x	P	x^*
5	0.25	1	8	0.188	2	12	0.156	3	17	0.109	4
6	0.125	0	9	0.125	1	13	0.094	2	18	0.078	3
7	0		10	0.062	0	14	0.062	1	19	0.049	2
			11	0		15	0.031	0	20	0.031	1
						16	0		21	0.016	0
									22	0	

$n=7$			$n=8$			$n=9$			$n=10$		
x	P	x^*	x	P	x^*	x	P	x^*	x	P	x^*
22	0.109	6	27	0.125	9	34	0.102	11	40	0.116	15
23	0.078	5	28	0.098	8	35	0.082	10	41	0.097	14
24	0.055	4	29	0.074	7	36	0.064	9	42	0.08	13
25	0.039	3	30	0.055	6	37	0.049	8	43	0.065	12
26	0.023	2	31	0.039	5	38	0.037	7	44	0.053	11
27	0.016	1	32	0.027	4	39	0.027	6	45	0.042	10
28	0.008	0	33	0.020	3	40	0.020	5	46	0.032	9
			34	0.012	2	41	0.014	4	47	0.024	8
			35	0.008	1	42	0.01	3	48	0.019	7
									49	0.014	6
									50	0.01	5

تابع - جدول (7) ولكيكسون

$P(T^+ \geq x) = p_r(T^+ \leq x^*)$

$n = 11$			$n = 12$			$n = 13$			$n = 14$		
x	P	x^*	x	P	x^*	x	P	x^*	x	P	x^*
48	0.103	18	56	0.102	22	64	0.108	27	73	0.108	32
49	0.087	17	57	0.088	21	65	0.095	26	74	0.097	31
50	0.074	16	58	0.076	20	66	0.084	25	75	0.086	30
51	0.062	15	59	0.065	19	67	0.073	24	76	0.077	29
52	0.051	14	60	0.055	18	68	0.064	23	77	0.068	28
53	0.042	13	61	0.046	17	69	0.055	22	78	0.059	27
54	0.034	12	62	0.039	16	70	0.047	21	79	0.052	26
55	0.027	11	63	0.032	15	71	0.040	20	80	0.045	25
56	0.021	10	64	0.026	14	72	0.034	19	81	0.039	24
57	0.016	9	65	0.021	13	73	0.029	18	82	0.034	23
58	0.012	8	66	0.017	12	74	0.024	17	83	0.029	22
59	0.009	7	67	0.013	11	75	0.020	16	84	0.025	21
			68	0.010	10	76	0.016	15	85	0.021	20
						77	0.013	14	86	0.018	19
						78	0.011	13	87	0.015	18
						79	0.009	12	88	0.012	17
									89	0.010	16

$n = 15$											
x	P	x^*	x	P	x^*	x	P	x^*	x	P	x^*
83	0.104	37	88	0.060	32	93	0.032	27	98	0.015	22
84	0.094	36	89	0.053	31	94	0.028	26	99	0.013	21
85	0.084	35	90	0.047	30	95	0.024	25	100	0.011	20
86	0.076	34	91	0.052	29	96	0.021	24	101	0.009	19
87	0.068	33	92	0.036	28	97	0.018	23			

جدول (8) مان وتني ذو اتجاهين

Critical Value of the Mann-Whitney U (Two − Tailed Testing)

n_2	α	n_1																	
		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	−	0	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
	.01	−	0	0	0	0	0	0	0	0	1	1	1	2	2	2	2	3	3
4	.05	−	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
	.01	−	−	0	0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	.05	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
	.01	−	−	0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6	.05	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
	.01	−	0	1	2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7	.05	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	37
	.01	−	0	1	3	4	6	7	9	10	12	13	15	16	18	19	21	22	27
8	.05	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
	.01	−	1	2	4	6	7	9	11	13	15	17	18	20	22	24	26	28	30
9	.05	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
	.01	0	1	3	5	7	9	11	13	16	18	20	22	24	27	29	31	33	36
10	.05	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
	.01	0	2	4	6	9	11	13	16	18	21	24	26	29	31	34	37	39	42
11	.05	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
	.01	0	2	5	7	10	13	16	18	21	24	27	30	33	36	39	42	45	48
12	.05	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
	.01	1	3	6	9	12	15	18	21	24	27	31	34	37	41	44	47	51	54
13	.05	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
	.01	1	3	7	10	13	17	20	24	27	31	34	38	42	45	49	53	56	60
14	.05	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74	78	83
	.01	1	4	7	11	15	18	22	26	30	34	38	42	46	50	54	58	63	67
15	.05	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
	.01	2	5	8	12	16	20	24	29	33	37	42	46	51	55	60	64	69	73
16	.05	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
	.01	2	5	9	13	18	22	27	31	36	41	45	50	55	60	65	70	74	79
17	.05	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93	99	105
	.01	2	6	10	15	19	24	29	34	39	44	49	54	60	65	70	75	81	86
18	.05	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
	.01	2	6	11	16	21	26	31	37	42	47	53	58	64	70	75	81	87	92
19	.05	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
	.01	3	7	12	17	22	28	33	49	45	51	56	63	69	74	81	87	93	99
20	.05	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127
	.01	3	8	13	18	24	30	36	42	48	54	60	67	73	79	86	92	99	105

جدول (8) مان وتني ذو اتجاه واحد

Critical Value of the Mann-Whitney U (One − Tailed Testing)

n_2	α	n_1																	
		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	0	0	1	2	2	3	4	4	5	5	6	7	7	8	9	9	10	11
	.01	−	0	0	0	0	0	1	1	1	2	2	2	3	3	4	4	4	5
4	.05	0	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
	.01	−	−	0	1	1	2	3	3	4	5	5	6	7	7	8	9	9	10
5	.05	1	2	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
	.01	−	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6	.05	2	3	5	7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
	.01	−	1	2	3	4	6	7	8	9	11	12	13	15	16	18	19	20	22
7	.05	2	4	6	8	11	13	15	17	19	21	24	26	28	30	31	35	37	39
	.01	0	1	3	4	6	7	9	11	12	14	16	17	19	21	23	24	26	28
8	.05	3	5	8	10	13	15	18	20	23	26	28	31	33	36	39	41	44	27
	.01	0	2	4	6	7	9	11	13	15	17	20	22	24	26	28	30	32	34
9	.05	4	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
	.01	1	3	5	7	9	11	14	16	18	21	23	26	28	31	33	36	38	40
10	.05	4	7	11	14	17	20	24	27	31	34	39	41	44	48	51	55	58	62
	.01	1	3	6	8	11	13	16	19	22	24	27	30	33	36	83	41	44	47
11	.05	5	8	12	16	19	23	27	31	34	38	42	46	50	54	57	61	65	69
	.01	1	4	7	9	12	15	18	22	25	28	31	34	37	41	44	47	50	53
12	.05	5	9	13	17	21	26	30	34	38	42	47	51	55	60	64	68	72	77
	.01	2	5	8	11	14	17	21	24	28	31	35	38	42	46	79	53	56	60
13	.05	6	10	15	16	24	28	33	37	42	47	51	56	61	65	70	75	80	84
	.01	2	5	9	12	16	20	23	27	31	35	39	43	47	51	55	59	63	67
14	.05	7	11	16	21	26	31	36	41	47	51	56	61	66	71	77	82	87	92
	.01	2	6	10	13	17	22	26	30	34	38	43	47	51	56	60	65	69	73
15	.05	7	12	18	23	28	33	39	44	50	55	61	66	72	77	83	88	94	100
	.01	3	7	11	15	19	24	28	33	37	42	74	51	56	61	66	70	75	80
16	.05	8	14	19	25	30	36	42	48	54	60	65	71	77	83	89	95	101	107
	.01	3	7	12	16	21	26	31	36	41	46	51	56	61	66	71	76	82	87
17	.05	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96	102	109	115
	.01	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77	82	88	93
18	.05	9	16	22	28	35	21	28	55	61	68	45	82	88	95	102	109	116	123
	.01	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88	94	100
19	.05	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	130
	.01	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	107
20	.05	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138
	.01	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

المراجع

1- الاحصاء التطبيقي، د. سمير سليم العبيدي، د. جمال ابراهيم البياني، دار شموع الثقافة، الزاوية، ليبيا.

2- الاحصاء التطبيقي، جامعة القدس المفتوحة، فلسطين.

3- الاحصاء التطبيقي، د. محمد زايد الدسوقي، مصر.

4- **Ronald E. Walpole , Raymond H. Myers, Sharon L. Myers, Probability and Statistics for Engineers and Scientists Eighth Edition.**

5- **Robert V. Hogg, Allen T. Craig, Introduction to Mathematical Statistics, Fifth Edition.**

6- **Murray R. Spiegel, John Schiller, R. Alu Srinivasan, Probability and Statistics, Third Edition.**